

**UCLA**

**UCLA Electronic Theses and Dissertations**

**Title**

Mobility Agent and Network Modeling for Decision Support in Transportation Systems

**Permalink**

<https://escholarship.org/uc/item/5qn8t6h1>

**Author**

Jiang, Qinhua

**Publication Date**

2024

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA  
Los Angeles

Mobility Agent and Network Modeling for Decision Support in Transportation Systems

A dissertation submitted in partial satisfaction  
of the requirements for the degree  
Doctor of Philosophy in Civil Engineering

by

Qinhua Jiang

2024



© Copyright by

Qinhua Jiang

2024

## ABSTRACT OF THE DISSERTATION

Mobility Agent and Network Modeling for Decision Support in Transportation Systems

by

Qinhua Jiang

Doctor of Philosophy in Civil Engineering

University of California, Los Angeles, 2024

Professor Jiaqi Ma, Chair

Transportation systems are complex systems encompassing interacting components such as infrastructure, vehicles, and travelers. The collective interconnection between these components makes decision-making processes within the system extremely challenging. While modeling the interaction between these components is crucial for effective transportation planning and traffic management, conventional transportation system models often struggle with limitations in adaptability, transferability, and accurate prediction.

This research starts with introducing a comprehensive agent-based modeling framework to evaluate the impact of new mobility options such as advanced vehicular technologies, evolving mobility patterns, and emerging vehicle usage behaviors, on transportation systems. The models developed within this framework provide a synthetic environment to evaluate the future impact on large-scale transportation systems from both demand and supply perspectives across various use cases, bridging existing gaps in adapting transportation systems to upcoming mobility innovations.

Aiming to address the limitations of existing agent-based transportation system modeling, especially the time-consuming and resource-intensive approaches in current human mobility pattern modeling, I present a state-of-the-art Artificial Intelligence (AI)- driven human mobility pattern synthesis model framework. This model employs a novel generative deep-learning approach for human mobility modeling and synthesis, using ubiquitous and open-

source data. Additionally, the model can be fine-tuned with local data, enabling transferable and accurate representations of mobility patterns across different regions.

The final segment of the research emphasizes the deployment of mobility AI network modeling in the real-world environment, especially the predictability for non-recurrent traffic conditions under complicated external environments. I present two deep-learning approaches for traffic state prediction in non-recurrent road conditions across large-scale networks and varying prediction time scales. The proposed traffic prediction models demonstrate superior performance, as validated by real-world data.

In essence, this thesis provides significant contributions to the domain of transportation system modeling by improving adaptability, transferability, and predictability in response to evolving mobility challenges. The developed tools and algorithms pave the way for the broader, real-world implementation of intelligent agent and network modeling in transportation system research.

The dissertation of Qinhua Jiang is approved.

Yifang Zhu

Adam S. Millard-Ball

Tierra Suzan Bills

Yueshuai Brian He

Jiaqi Ma, Committee Chair

University of California, Los Angeles

2024

*To my wife, Xiaohan, my daughter, Harper, and my parents, Heyang and  
Juxiang*

# Contents

<b>Abstract</b>	<b>ii</b>
<b>List of Figures</b>	<b>x</b>
<b>Acknowledgements</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation	1
1.2 Objective and System Framework	3
1.2.1 Part I: Agent-Based Transportation System Modeling	5
1.2.2 Part II: Mobility AI Agent Modeling	6
1.2.3 Part III: Mobility AI Network Modeling	6
1.2.4 Interaction Between Mobility Agents and Network	7
1.2.5 Interaction Between the Digital Representation and the Physical Transportation System	8
1.3 Contributions	9
<b>2 Evaluating the impact of connected and automated vehicles on transportation system performance</b>	<b>11</b>
2.1 Introduction	12
2.2 Related Work	15
2.3 Methodology	18
2.3.1 Integration of CAV-ABM and Assignment Model	18
CAV-ABM	18
Network Capacity Adjustment	19
Multiclass User Equilibrium Highway Assignment	21
2.3.2 Scenario Design	24
CAV Base Case Scenario	24
Demand Management Policy Scenarios	25
2.3.3 Performance Metrics	28
System Impact Metrics	28
Equity Metrics	29
2.4 Results and Discussion	31
2.4.1 Transportation System Impacts Analysis	31
2.4.2 Results and Discussion	32

2.4.3	Travel Equity Analysis . . . . .	36
2.5	Conclusion . . . . .	40
<b>3</b>	<b>Evaluating the traffic and environmental impact of telework and remote services on transportation systems</b>	<b>42</b>
3.1	Introduction . . . . .	43
3.2	Related Work . . . . .	46
3.3	Methodology . . . . .	48
3.3.1	Tele-X Survey . . . . .	49
3.3.2	Activity-Based Travel Demand Modeling . . . . .	52
	Telework Model . . . . .	53
	Teleservice Model . . . . .	56
	Agent-Based Transportation Simulation . . . . .	59
	Emission Factor (EMFAC) Model . . . . .	62
	Scenario Design . . . . .	63
3.4	Results and Discussion . . . . .	65
3.4.1	Network Experiment Results . . . . .	65
	Activity Analysis . . . . .	65
	Travel Analysis . . . . .	66
	Emission Analysis . . . . .	69
3.4.2	Corridor Experiment Results . . . . .	70
3.5	Conclusion . . . . .	75
<b>4</b>	<b>Gap assessment and demand projection of public charging infrastructure in electrified transportation systems</b>	<b>77</b>
4.1	Introduction . . . . .	78
4.1.1	Related Work . . . . .	79
4.1.2	Main Contribution . . . . .	82
4.2	Methodology . . . . .	84
4.2.1	Model Overview . . . . .	84
4.2.2	Travel Profile Generation . . . . .	85
	Activity-Based Transportation Simulation System . . . . .	85
	Scenario-Specific Travel Demand and Transportation Network . . . . .	86
4.2.3	Charging Profile Estimation . . . . .	89
	Future Electrification Levels . . . . .	89
	Initial SOC . . . . .	90
	Charging Rule Definition . . . . .	90
4.2.4	Demand Aggregation and Equity Assessment . . . . .	93
	Charging Demand Aggregation and Charger Estimation . . . . .	93
	Charging Equity Measurements . . . . .	96
4.3	Case Study and Data Specification . . . . .	100
4.3.1	Travel Data Specification . . . . .	100
4.3.2	Electrification and Charging Data Specification . . . . .	102
	Electrification Data . . . . .	102

	Charging Data . . . . .	105
4.4	Results and Discussion . . . . .	107
4.4.1	Network-Level Charging Demand Analysis . . . . .	107
	Network-level Temporal Pattern Analysis . . . . .	108
	Network-level Spatial Pattern Analysis . . . . .	110
4.4.2	Zone-level Charging Demand Variation Across Scenarios . . . . .	111
4.4.3	Zone-level Charging Demand Variation Across Functional Areas . . . . .	114
4.4.4	Estimation of Public Chargers and Installation Cost for the Year 2035 in LA County . . . . .	115
4.4.5	Equity Assessment of Public Charging . . . . .	119
4.5	Conclusion . . . . .	122
<b>5</b>	<b>Deep activity model: a generative approach for human mobility pattern synthesis</b> . . . . .	<b>125</b>
5.1	Introduction . . . . .	126
5.2	Related work . . . . .	130
5.3	Problem Formulation . . . . .	132
5.4	Dataset . . . . .	133
5.4.1	Household Travel Survey (HTS) . . . . .	133
5.4.2	SCAG ABM Data . . . . .	134
5.4.3	Data Preparation . . . . .	134
5.5	Methodology . . . . .	136
5.5.1	Model Architectures . . . . .	137
5.5.2	Loss functions . . . . .	139
5.5.3	Data Balancing . . . . .	140
5.5.4	Model Transfer . . . . .	142
5.5.5	Activity Location Assignment and Network Traffic Loading . . . . .	143
5.6	Experiments and Results . . . . .	147
5.6.1	Training and Evaluation Methods . . . . .	147
5.6.2	Baseline Models . . . . .	149
5.6.3	Evaluation on Activity Generation . . . . .	150
	Distribution Similarity . . . . .	150
	Loss Term Ablation Study . . . . .	152
	Contextual Variation . . . . .	153
	Interdependency among Household members and Activity in Activity Generation . . . . .	155
	Model Transferability . . . . .	156
	Data Balancing Algorithm Evaluation . . . . .	159
5.6.4	Validation of ALA and Network Traffic Loading . . . . .	159
5.6.5	The Influence of Model Complexity and Data Size . . . . .	163
5.7	Conclusion and Future Work . . . . .	165
<b>6</b>	<b>Deploying scalable traffic prediction models in real-world large trans- portation networks during hurricane evacuations</b> . . . . .	<b>167</b>



6.1	Introduction . . . . .	168
6.2	Related Work . . . . .	172
6.2.1	Model-based evacuation traffic modeling . . . . .	172
6.2.2	Data-driven evacuation traffic modeling . . . . .	172
6.3	Methods . . . . .	174
6.3.1	Methodology Overview . . . . .	174
6.3.2	Long-term Congestion State Prediction Model . . . . .	176
6.3.3	Short-term Traffic Speed Prediction Model . . . . .	179
6.4	Experiment Design . . . . .	182
6.4.1	Data Preparation . . . . .	182
6.4.2	Baseline Models . . . . .	184
6.4.3	Evaluation Metrics . . . . .	188
6.5	Results . . . . .	190
6.5.1	Long-term congestion state prediction . . . . .	190
6.5.2	Short-term traffic speed prediction . . . . .	194
6.6	Conclusion . . . . .	197
<b>7</b>	<b>Leveraging data-centric AI for work zone traffic impact prediction</b>	<b>199</b>
7.1	Introduction . . . . .	200
7.2	Data Integration and Reconstruction . . . . .	204
7.2.1	Multi-Context Work Zone Data . . . . .	204
7.2.2	Data Integration and Space-Time Traffic Image Generation . . . . .	205
7.3	Methodology . . . . .	208
7.3.1	Problem Definition . . . . .	208
7.3.2	Model Structure . . . . .	208
	Model Overview . . . . .	208
	Encoder Layers . . . . .	211
	Attention Layer . . . . .	211
	Decoder Layers . . . . .	213
7.3.3	Loss Function . . . . .	214
7.4	Experimentation . . . . .	214
7.4.1	Performance Metrics . . . . .	214
7.4.2	Baseline Models . . . . .	216
7.5	Results and Performance Evaluation . . . . .	217
7.5.1	System-Level Performance Analysis . . . . .	217
7.5.2	Event-Level Spatial-Temporal Performance Analysis . . . . .	220
7.6	Conclusion . . . . .	222
<b>8</b>	<b>Conclusions and Future Work</b>	<b>223</b>
8.1	Conclusions . . . . .	223
8.2	Future Work . . . . .	225
	<b>Bibliography</b>	<b>260</b>

# List of Figures

1.1	System framework of the intelligent agent and network modeling for transportation systems . . . . .	4
2.1	System design of SCAG ABM . . . . .	20
2.2	Model road network in Southern California . . . . .	22
2.3	Model Integration Framework . . . . .	24
2.4	Ratio of mean speed to free-flow speed during AM peak: (a) SCAG; (b) Zoom in view of (a); (c) CAV; and (d) Zoom in view of (c). . . . .	33
2.5	Ratio of mean speed to free-flow speed during PM peak: (a) SCAG model; (b) Zoom in view of (a); (c) CAV base model; and (d) Zoom in view of (c). . . . .	34
2.6	Comparison of mode share between SCAG, CAV demand-only, and CAV demand and supply. . . . .	35
2.7	Equity Index across Scenarios for (a) Number of Trips per Household and (b) Travel Distance per Household . . . . .	38
2.8	Equity Index across Scenarios on Travel Accessibility . . . . .	39
3.1	Methodology framework . . . . .	49
3.2	Modeling framework of the LA-Sim model . . . . .	60
3.3	Percentage change of (a) VMT (b) NOx (c) PM 2.5 in current baseline scenario compared to the Pre-pandemic scenario (Scenario A) . . . . .	68
3.4	Spatial distribution of non-work trip origins from elsewhere, ending in Westwood communities . . . . .	71
3.5	Temporal distributions of (a) traffic volume and (b) travel speed on I-405 and (c) traffic volume and (d) travel speed on Westwood segment . . . . .	72
3.6	NOx and PM 2.5 percentage change relative to Pre-pandemic scenario on I-405 for Westwood . . . . .	74
4.1	Model Framework Overview . . . . .	84
4.2	Initial SOC definition . . . . .	89
4.3	The public charging decision-making process . . . . .	92
4.4	Relation between charging demand and required chargers . . . . .	95
4.5	Major Steps in the Charging Demand Prediction Model . . . . .	99
4.6	(a) Census tracts of LA county (CTs); (b) The road transportation network . . . . .	100
4.7	Travel demand temporal distribution by trip purposes: a) All trips, b) Work trips, and c) Non-work trips . . . . .	101

4.8	2035 EV adoption rate across CTs under (a) moderate electrification level and (b) high electrification level . . . . .	106
4.9	Temporal distribution of charging demands from weekday scenarios and weekend scenarios of (a) High electrification plus low initial SOC; (b) High electrification plus high initial SOC; (c) Moderate electrification plus low initial SOC; and (d) Moderate electrification plus high initial SOC . . . . .	109
4.10	LA county (a) Geographic regions (Wikimedia, 2022); and (b) Population density distribution (California Air Resources Board (CARB), 2022) . . . . .	110
4.11	Spatial distribution of peak charging demand from weekday scenarios of (a) High electrification plus low initial SOC; (b) High electrification plus high initial SOC; (c) Moderate electrification plus low initial SOC; and (d) Moderate electrification plus high initial SOC . . . . .	112
4.12	Charging demand variation across weekday scenarios: (a) Distribution of mean and standard deviation of normalized charging demand; and (b) Distribution of standard deviation divided by the mean of normalized charging demand . . . . .	113
4.13	Temporal charging demand distribution of a Test CT across all network scenarios. a) The location of the test CT. b) 24-hour charging demand distribution of the test CT. The semi-transparent bands represent the upper and lower bound of the charging demand across network scenarios . . . . .	114
4.14	Spatial distribution of charging demand with all scenarios under high electrification level plus low initial SOC y at (a) Working area; (b) School area; (c) Shopping area; and (d) Entertaining area . . . . .	116
4.15	Temporal distribution of charging demand and estimated chargers for a single zone under a specific simulation scenario . . . . .	117
4.16	Charger access probability across different socio-demographic groups for 2022 (a, d, g), 2035 moderate electrification level (b, e, h), and 2035 high electrification level (c, f, i) . . . . .	120
4.17	Charger access disparity of 2022 and 2035 across different socio-demographic groups: (a) DAC/non-DAC groups; (b) income groups; (c) ethnicity groups . . . . .	121
5.1	Model human mobility pattern using household travel survey (HTS) data. (a) HTS data includes information about each household member’s social-demographics, the household itself, and their daily non-commercial travel. This data covers daily non-commercial travel by all modes, along with details about the travelers, their households, and their vehicles [163]. (b) Typical weekday and weekend activity chains in HTS. . . . .	128
5.2	Workflow of activity chain generation . . . . .	132
5.3	Deep Activity model architecture. (a) Input data construction. (b) Transformer-based network architecture with well-designed data injection. . . . .	137
5.4	Distributions of activity distances and angular deviations across sub-regions in LA . . . . .	145
5.5	Detailed analysis and comparison of activity generation on (a)(b) temporal dynamics, (c) activity chain length, (d) activity duration, and (e) activity type distribution. . . . .	149

5.6	Activity patterns in weekdays and weekends . . . . .	154
5.7	Activity patterns across age groups . . . . .	155
5.8	Attention weights reveals the interdependency among household members and activities. . . . .	156
5.9	Distribution comparison for datasets from three regions, showing significant different activity patterns. Activity type labels are excluded because dataset of Mexico City only contain 10 types of activity that are different from CA and Puget Sound region. . . . .	158
5.10	Data balancing is performed on Mexico City. . . . .	160
5.11	LA County Map and Freeway Network. . . . .	161
5.12	Validation for ALA and traffic loading at network level. . . . .	162
5.13	Validation for ALA and traffic loading at corridor level. . . . .	163
5.14	Performance evaluation for transformer models with different complexity. . .	164
5.15	Training data size effect on decoder-only transformer and LSTM. . . . .	165
6.1	The framework of the multi-scale hurricane evacuation traffic prediction model	175
6.2	Model framework for long-term hurricane evacuation traffic prediction model	177
6.3	Model framework for short-term hurricane evacuation traffic prediction model	180
6.4	IMRCP system integration diagram . . . . .	182
6.5	Roadway network of Louisiana . . . . .	183
6.6	Example speed data during hurricane evacuation: (a) Link location; (b) 7-day speed plot of an example link during hurricane Ida . . . . .	185
6.7	Example speed data during hurricane evacuation: (a) Link location; (b) 7-day speed plot of an example link during hurricane Ida . . . . .	191
6.8	Comparison of prediction performance with LSTM and baseline models during hurricane Ida ((a), (b), (c): test link 1 with 1-, 3-, and 6-hour horizon; (d), (e), (f): test link 2 with 1-, 3-, and 6-hour horizons) . . . . .	195
7.1	Multi-context data integration pipeline . . . . .	204
7.2	Spatial distribution of work zone, detectors, and road segments in Maryland transportation network . . . . .	205
7.3	Work zone space-time traffic image generation: (a) 2D space-time traffic matrix; (b) Two examples of converted 2D space-time traffic images (work zone traffic speed and historical traffic speed) . . . . .	206
7.4	Model structure of AMCNN-ED . . . . .	209
7.5	Work zone space-time traffic image processing . . . . .	216
7.6	Examples of event-level speed prediction performance . . . . .	221

## Acknowledgements

First and foremost, I would like to express my gratitude to my advisor, Dr. Jiaqi Ma, for his unwavering support, guidance, and mentorship throughout my PhD journey. His expertise, patience, and encouragement have been invaluable, and I am incredibly grateful for the countless hours he has dedicated to helping me grow as a researcher and professional. Without his insight and direction, this dissertation would not have been possible.

I would also like to thank my labmates in the UCLA Mobility Lab for their support and collaboration. Working alongside such talented and dedicated individuals has been both inspiring and motivating. The discussions, feedback, and shared experiences have enriched my research and made this journey far more enjoyable.

Finally, my deepest appreciation goes to my family. To my parents, Heyang and Juxiang, who have always believed in me and supported my aspirations—I am endlessly grateful for your sacrifices and encouragement. To my wonderful wife, Xiaohan, whose love, patience, and understanding have been my greatest strength, especially during the most challenging times of this journey. And to my newborn daughter, Harper, whose arrival has brought immeasurable joy and renewed purpose to my life. This work is dedicated to you all.

## Curriculum Vitae

### Education

Beijing Jiaotong University	2012/09 - 2016/06
B.S. in Civil Engineering	
Beijing Jiaotong University	2016/09 - 2019/06
M.S. in Civil Engineering	
University of California, Los Angeles	2020/09 - present
Ph.D. student in Civil Engineering	

### Awards

UTC PSR Doctoral Student of the Year	2022/03
--------------------------------------	---------

### Publications

1. **Q. Jiang**, N. Zhang, B. Y. He, C. Lee, and J. Ma, "Large-scale public charging demand prediction with a scenario-and activity-based approach," *Transportation Research Part A: Policy and Practice*, vol. 179, pp. 103935, 2024.
2. **Q. Jiang**, B. Y. He, C. Lee, and J. Ma, "Deploying scalable traffic prediction models for efficient management in real-world large transportation networks during hurricane evacuations," *arXiv preprint arXiv:2406.12119*, 2024.
3. **Q. Jiang**, X. Liao, Y. Gong, and J. Ma, "An Attention-Based Multi-Context Convolutional Encoder-Decoder Neural Network for Work Zone Traffic Impact Prediction," *arXiv preprint arXiv:2405.21045*, 2024.
4. B. Y. He, **Q. Jiang**, H. Ma, and J. Ma, "Multi-agent multimodal transportation simulation for mega-cities: Application of Los Angeles," *Procedia Computer Science*, vol. 238, pp. 736–741, 2024.

5. X. Liao, **Q. Jiang**, B. Y. He, C. Kuai, and J. Ma, "Deep Activity Model: A Generative Approach for Human Mobility Pattern Synthesis," *arXiv preprint arXiv:2405.17468*, 2024.
6. D. Nian, **Q. Jiang**, and J. Ma, "Developing highway capacity manual capacity adjustment factors for connected and automated traffic at two-way stop-controlled intersections," *Journal of Transportation Engineering, Part A: Systems*, vol. 149, no. 2, pp. 04022144, 2023.
7. **Q. Jiang**, D. Nian, Y. Guo, M. Ahmed, G. Yang, and J. Ma, "Evaluating connected vehicle-based weather responsive management strategies using weather-sensitive microscopic simulation," *Journal of Intelligent Transportation Systems*, vol. 27, no. 1, pp. 92–110, 2023.
8. **Q. Jiang**, B. Y. He, Q. Yu, L. Li, and J. Ma, "Are Telework and Teleservices the Solution for Reducing Traffic Congestion and Emissions? Findings from a Comprehensive Activity-Based Modeling Analysis."
9. **Q. Jiang**, B. Y. He, and J. Ma, "Connected automated vehicle impacts in Southern California part-II: VMT, emissions, and equity," *Transportation Research Part D: Transport and Environment*, vol. 109, pp. 103381, 2022.
10. B. Y. He, **Q. Jiang**, and J. Ma, "Connected automated vehicle impacts in Southern California part-I: Travel behavior and demand analysis," *Transportation Research Part D: Transport and Environment*, vol. 109, pp. 103329, 2022.
11. **Q. Jiang**, B. Schroeder, J. Ma, L. Rodegerdts, B. Cesme, A. Bibeka, and A. Morgan, "Developing highway capacity manual capacity adjustment factors for connected and automated traffic on roundabouts," *Journal of Transportation Engineering, Part A: Systems*, vol. 148, no. 5, pp. 04022014, 2022.

# Chapter 1

## Introduction

### 1.1 Motivation

Transportation systems are intricate networks comprising various interconnected components: road infrastructure (e.g., roads, lanes, intersections, and traffic lights), vehicles (e.g., cars, buses, and trucks), and most importantly, humans (e.g., drivers, passengers, traffic controllers) with their highly sophisticated and unpredictable behavior [41]. Additionally, numerous external factors such as weather conditions, new mobility trends, social events, and traffic operations can significantly influence travel behavior and traffic patterns [205]. The interaction of these elements creates a large-scale, integrated, and complex system.

Transportation system models offer a structured representation of real-world transportation systems. Fundamentally, an integrated transportation system model that meets diverse planning needs includes both agent modeling that generates human travel activities and network modeling that assigns trips to the respective networks, generating traffic condition estimations accordingly [175]. These models are indispensable tools for evaluating the impacts of various factors and predicting the performance of transportation systems, which is crucial for effective decision-making for transportation system planning and management.

Over the past few decades, numerous advanced statistical and simulation models have



been developed to enhance the decision-making process in transportation system planning and traffic management [106]. These include activity-based modeling for travel demand generation, dynamic traffic assignments for network traffic modeling, and the integration of advanced travel demand and network simulation models, highlighting an increasing need for greater precision, dynamism, and granularity. However, the inherent complexity of transportation systems means that many components and their interactions cannot be fully captured by conventional model-driven demand generation or agent-based network simulations. Several challenges have emerged across the planning, deployment, and operation phases of transportation system modeling, such as:

- **Scenario Adaptability:** Current demand generation and network simulation models, whether trip-based or activity-based approaches, focus solely on existing mobility and network scenarios. They struggle to adapt to changing contexts, such as assessing impacts and predicting demand for scenarios involving new mobility options, evolving behavior patterns, or novel vehicle usage practices.
- **Regional Transferability:** Existing travel demand generation models heavily depend on expert knowledge and are implemented using hard-coded rules and heuristics. This makes their development extremely laborious, time-consuming, and resource-intensive. Moreover, these well-calibrated models are highly customized to the socio-demographics and network characteristics of a particular area, making it challenging to transfer them to regions with different socio-demographic and geospatial characteristics.
- **Traffic Predictability:** Although current transportation network models can simulate daily traffic variations with high spatial and temporal granularity for large roadway networks, they are often unable to capture the traffic patterns of low-frequency, non-recurrent events, such as evacuation-induced traffic or construction-caused local congestions. These limitations constrain current traffic models from supporting

decision-making and providing efficient management strategies during special roadway scenarios.

To address these challenges, advanced models and computational techniques are needed to better understand travel behavior and traffic flow and tackle the challenges of prediction, monitoring, management, and planning in transportation systems.

## 1.2 Objective and System Framework

This dissertation aims to address challenges in existing transportation system modeling by leveraging mobility agent and network modeling to develop a digital twin of the physical transportation system. This digital twin aims to accurately model the behavior of all system components, including agents (i.e., travelers and vehicles) and roadway networks. The proposed framework begins by extracting multi-source data from the physical transportation system, encompassing foundational, processed, and synthetic data. This curated data repository is fundamental for learning the inherent correlation between different transportation components and is used to construct a digital representation of the physical transportation system. The output of the transportation system digital twin will support decision-making in regional transportation planning policies and traffic management strategies. The implementation of these strategies will, in turn, impact the physical transportation system, iteratively influencing the construction and evolution of the digital twin in a feedback loop.

The core of this transportation system digital twin is its digital representation of the transportation system. The fundamental concept involves modeling agents and networks, as well as their interactions within the transportation system. This collaborative effort forms a digital simulacra of individual travelers and roadway infrastructure, enabling the digital representation of a dynamic and constantly evolving transportation system.

The construction of the digital representation of the transportation system includes three main components, as presented in Fig 1.1: agent-based transportation system modeling,

AI mobility agent modeling, and AI mobility network modeling. This dissertation first introduces agent-based transportation system modeling to enhance the adaptability of existing transportation system modeling approaches. This allows for accurate modeling from both demand and supply perspectives, considering the interactions between individual travelers and the network. Subsequently, AI techniques are integrated into the digital transportation system to develop next-generation AI agent modeling and AI network modeling, addressing transferability and predictability challenges encountered in current transportation system models. Specifically, the key components of the transportation system digital twin are presented as follows:

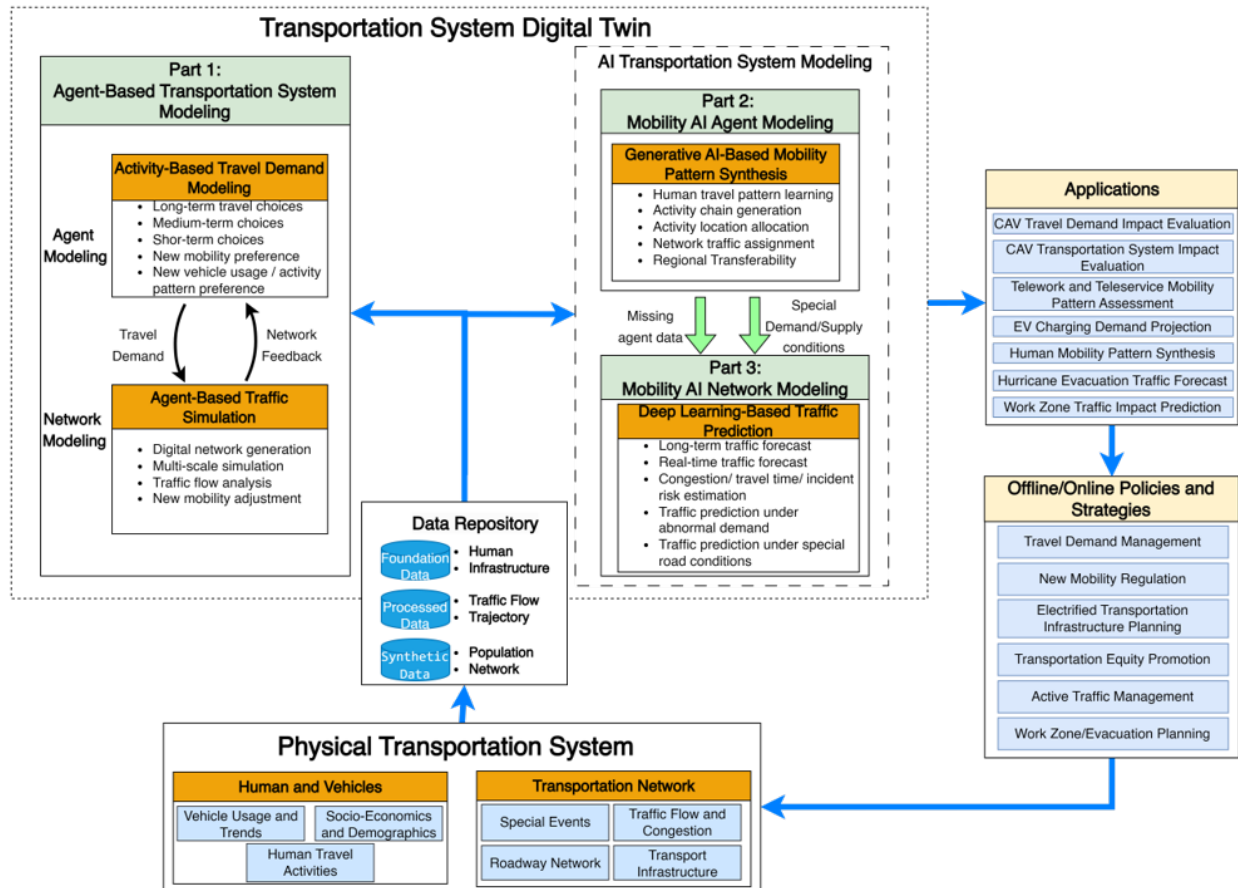


Figure 1.1: System framework of the intelligent agent and network modeling for transportation systems

## 1.2.1 Part I: Agent-Based Transportation System Modeling

Chapter 2 [99, 117], Chapter 3 [161], and Chapter 4 [122] form Part I of this dissertation. Part I lays the foundation for constructing a digital simulacra of a transportation system. It produces a digital twin of the transportation system by integrating agent behavior modeling and transportation network modeling. The agent behavior modeling adopts an activity-based travel demand modeling approach to simulate the individual decision-making processes during daily travel. This encompasses travel choices at various scales, from long-term decisions like home and work locations, to medium-term choices like daily activity frequencies, and short-term choices like travel mode selection for each trip. Furthermore, by incorporating additional data such as stated preference survey data, people’s travel preference on novel scenarios such as new mobility or new vehicle usage practice can be well captured in a digital simulacra of travel demand generation. The network modeling creates a digital replica of the roadway network in the simulation environment, representing the infrastructure serving agents’ travel. It utilizes agent-based traffic simulation techniques to generate a detailed digital replica of people’s travel trajectories within the network. The network is also flexible to adjust for any infrastructure or accessibility upgrade in response to the era of future mobility options. The agent and network modeling serve collaboratively to create a digital representation of the transportation system with high spatial and temporal resolution, forming realistic travel demand and traffic flow patterns on the transportation network.

Importantly, this integrated modeling structure of mobility agents and networks enhances the capability of existing models to address new mobility scenarios, such as connected and autonomous vehicles (CAVs), new vehicle usage patterns like electric vehicles (EVs), and emerging travel behaviors like teleworking or remote services. This approach allows existing transportation system models to make offline adaptations to new mobility challenges with minimal model redevelopment effort, generating decision support for long-term planning strategies of the transportation system.

## 1.2.2 Part II: Mobility AI Agent Modeling

Chapter 5 [147] constitutes Part II of this research. In Part II, we further enhance the capability of agent modeling based on the structure developed in Part I. We adopt generative AI concepts to model agents and their interaction with the network in the transportation system. This AI agent modeling marks a substantial advancement in simulating human mobility behavior within our digital twin and can function as the next generation of travel demand models. To generate travel demand, AI agents first learn the activity patterns associated with various identities, such as work, school, or leisure. They then develop travel chains based on learned preferences, including preferred activity types, frequencies, trip routes, and transportation modes, which adjust based on past experiences and new circumstances. These learned travel choices by the AI agents will, in turn, be reflected in the regional daily travel demand generation, consequently forming the digital replica of mobility patterns in the mobility network.

The transferability of a pre-trained AI agent model highlights the most significant enhancement as opposed to conventional travel demand models. The AI agent modeling allows for seamless behavior adaptation across different regions and cultures, making it versatile for various transportation systems. This transferability is achieved through fine-tuning techniques. This allows for training the base model initially on a large, diverse dataset and then adapting it to specific regions or cultures by training on smaller, localized datasets. This approach enables the model to retain general mobility patterns while learning region-specific behaviors, facilitating efficient deployment across different urban environments with minimal additional training.

## 1.2.3 Part III: Mobility AI Network Modeling

Chapter 6 [119] and Chapter 7 [121] form Part III of this dissertation. Part III enhances the capability of AI agent and network modeling under special conditions by introducing AI

traffic modeling. This serves as a supplement to AI agent and network modeling in extreme scenarios where agent data is completely missing, or special demand or supply side conditions arise, such as emergency evacuations or temporary road construction-induced traffic patterns. AI traffic modeling can predict traffic conditions by considering multiple factors, such as weather and special events, to accurately forecast future traffic states. The traffic state estimation capabilities provide both long-term and real-time traffic forecast insights into mobility system conditions, including areas with limited data due to signal loss, device malfunction, or low sensor coverage. This prediction goes beyond traffic congestion and can further provide in-advance estimation for congestion, travel time, and incident likelihood of the target roadway network. The AI network model also aims to understand underlying causes and potential chain effects, allowing for interaction with dynamic demand modeling.

Additionally, the evolution capability of AI traffic modeling enables adaptation to long-term changes in urban roadway, policy, and human behavior, effectively evolving with the region they represent. Consequently, the AI network can adjust its model over time to accommodate significant changes, such as new infrastructure developments, shifts in population density, or evolving transportation policies.

#### **1.2.4 Interaction Between Mobility Agents and Network**

The mobility system digital twin developed in this dissertation allows for a bidirectional analysis of how individual travelers (i.e., agents) affect transportation network performance and how variations in the network interactively impact agents' decisions. By integrating intelligent mobility agent modeling with intelligent network modeling, a digital representation of a real-world mobility system is established, forming an interactive digital ecosystem. This integration facilitates sophisticated simulations where individual agent behaviors both affect and are affected by network conditions. For instance, the aggregate behavior of agents forms the network's spatial-temporal pattern and evolution, while the network's traffic state

interactively influences agents' travel behavior in terms of route choices, mode selection, and anticipated travel times. This reversible interaction enables the digital twin to capture complex scenarios, such as temporary demand surges, emergent traffic patterns, the impact of policy changes, and the collective effects of local disturbances across the entire mobility system. Additionally, this integration of intelligent agents and networks allows for flexible scenario experiments. By setting up simulated changes to either agent behaviors or network conditions, we can generate system-wide outcomes for specific scenarios, providing valuable insights for transportation system planning and traffic management strategies.

### **1.2.5 Interaction Between the Digital Representation and the Physical Transportation System**

The digital representation of the transportation system, built on intelligent agent and network modeling, offers numerous applications across various sectors in transportation planning and management. Specifically, this dissertation emphasizes the model development and analysis of transportation system digital twins for use cases such as:

- CAV impact evaluation on travel demand and transportation system performance,
- Assessment of telework and teleservices on human mobility patterns,
- EV charging demand projection in electrified transportation systems,
- Human mobility pattern synthesis and evaluation,
- Traffic flow prediction during hurricane evacuations, and
- Traffic impact prediction during work zone activities

Different modules in the digital twin collaboratively provide both offline and online decision support for the physical transportation system. On one hand, agent-based transportation

system modeling within the digital twin framework develops robust statistical models based on archived data to predict and analyze human travel patterns and traffic dynamics in the target region. This offers valuable insights for urban planners and policymakers to evaluate the impact of future infrastructural changes, policy implementations, and technological advancements. It supports decision-making processes aimed at enhancing transportation efficiency, reducing emissions, and improving overall urban mobility in densely populated areas. On the other hand, mobility AI agent and network modeling, utilizing pre-trained models and fine-tuning techniques, opens the possibility to harness not only historical data but also more recent mobility data in modeling agent travel behavior and network conditions. This allows for short-term or even real-time responses to roadway conditions or travel patterns, such as temporary demand changes due to special events or emergent land closures caused by incidents or construction. This enables accurate and timely predictions for travel and traffic patterns, significantly supporting decision-making for immediate responses to behavioral and infrastructural variations.

### 1.3 Contributions

This dissertation enhances decision-making in transportation systems by leveraging intelligent mobility agent and network modeling to create a digital representation of the physical transportation system. This representation enables interaction between the physical transportation system and the digital twin, providing both online and offline decision support under sophisticated human mobility and network dynamics. The key contributions of this dissertation are summarized as follows:

- **Development of a digital twin framework for intelligent decision support in transportation systems.** This dissertation develops a data-centric transportation system digital twin framework to support both online and offline decision-making. By incorporating massive multi-context data from the physical transportation system,



along with mobility AI agent and network modeling, the limitations of existing models in terms of adaptability, transferability, and predictability are significantly enhanced.

- **Creation of an agent-based transportation system modeling approach to enhance model adaptability.** Part I of the dissertation develops a series of model integration frameworks that facilitate the incorporation of agent-based modeling with additional multi-source data from both demand and supply perspectives. Compared with conventional agent-based modeling approaches, the proposed models are tailored to adapt existing travel demand and network simulation models to new mobility options, emerging travel behaviors, and novel vehicle usage patterns.
- **Design of a mobility AI agent modeling architecture to facilitate model transferability.** Part II of the dissertation explores the potential of a fully data-driven AI agent modeling approach for synthesizing human mobility patterns, aiming to automate the resource-intensive and laborious process associated with traditional travel demand and network simulation models. The model is the first of its kind, evaluated in an agent-based traffic assignment environment, and it demonstrates significant fidelity and efficiency compared to conventional travel demand and traffic assignment models.
- **Development of a mobility AI network modeling approach to promote traffic state predictability.** Part III of the dissertation focuses on the development and deployment of AI network traffic prediction models to handle non-recurrent roadway and traffic scenarios. The models involve the incorporation of multi-context data and novel machine-learning structures, facilitating accurate spatial-temporal traffic predictions during special scenarios and enhancing the predictability of network traffic models for both regional transportation planning and real-time traffic management.

## Chapter 2

# Evaluating the impact of connected and automated vehicles on transportation system performance

Connected and automated vehicle (CAV) technologies are likely to have significant impacts on people's travel behaviors and the performance of transportation systems. This study investigates the impacts of CAVs from various aspects, including vehicle miles traveled (VMT), emissions, and transportation equity in Southern California. A comprehensive model is developed by incorporating the supply-side improvement of CAVs, a modified activity-based demand model supported by survey data, and a multi-class highway assignment model. The simulation results showed that VMT and emissions would increase by 10%, and CAVs could worsen travel equity across income groups. To reduce the negative impacts caused by CAVs, we proposed and evaluated a series of travel demand management policies. The results indicated that all policies help to reduce the VMT and emission growth, while their performances in enhancing travel equity vary across metrics including accessibility, travel frequency, and travel distance.

## 2.1 Introduction

Connected and automated vehicle (CAV) technologies are one of the most rapidly developing automobile technologies [276]. CAV technologies are the integration of connected vehicle (CV) and automated vehicle (AV) technologies. CVs can enable communications between vehicles, mobile devices, and infrastructure, thus providing potential benefits for both the drivers and transportation systems [130]. Vehicles that can be defined as AVs must have automated control of at least some of a safety-critical control function (e.g., steering or braking) [199]. Although both CV and AV technologies have gained rapid growth along their independent paths, they should be bonded more strongly to contribute to the mobility and safety of the transportation system [240]. There will be significant implications on different aspects of transportation, which cover a range from the single vehicle level [134] to the transportation system level [169]. As widely known, there are numerous benefits of CAVs, including enhanced driver/passenger safety [134], increased roadway capacity [4], reduced congestion [169], and potential reduction in emissions [149]. CAVs can eliminate the possibility of human error, which causes 94% of accidents [188]. In addition, with the help of in-vehicle equipment and connection with infrastructure, the drivers' perception-reaction time becomes shorter, and the lateral and longitudinal distance between vehicles can be significantly reduced. Finally, applications of CAVs such as platooning can decrease congestion and thereby improve fuel economy. CAV technology also disproportionately benefits younger, older, and disabled people by enhancing their accessibility to transportation options currently out of reach [58].

Since still at an early stage, the rapid development of CAV technologies has caused many problems from not only the vehicle operation side but also the integration into the existing transportation system. From the vehicle level, some of these challenges include system failure, interactions with conventional vehicles, and inappropriate operation errors to unexpected circumstances [73]. From the perspective of social economics, the high cost of the early stage CAVs compared with conventional vehicles may become a downside effect on

individuals' travel equity [17]. Ultimately, it's significant to investigate to what extent CAVs impact transportation systems in the context of induced travel demand and enhanced network capacity. For instance, will the increased demand offset the benefits brought by CAVs? How well can the increased roadway capacity help alleviate the strain on the network with CAV deployment? It is necessary to further strengthen the potential benefits of CAVs and eliminate concerns that might become a drag on the future development of this technology so that CAVs can be widely accepted in the future.

This study adopted an activity-based approach to evaluate the comprehensive impacts of CAV technology on the transportation system in Southern California. Considering the scope of this study, we divided the whole study into two consecutive papers: Part one and Part two. Part one [99] explored the changes in travel behavior and predicted travel demand that fundamentally incorporates CAV-related changes and addressed the following problems:

- Collected Southern California residents' willingness to use CAVs for daily travel, including both the personally owned CAVs (PAV) and shared CAVs (SAV).
- Captured the long-term to short-term changes in CAV users' travel behaviors.
- Incorporated people's CAV-related behavior changes in a state-of-the-art activity-based modeling (ABM) framework and predicted the new travel demand and trip patterns.

This chapter is the continued study of Part one and concentrates on incorporating the enhancement on the supply side by infrastructure-enabled automation and evaluating the impacts of CAV technology and policies on the transportation system and the associated equity concerns. The contributions of the study are:

- This study captures actual behavior preferences using stated preference survey data and models these changes through advanced activity choice models, rather than relying on assumed behavior patterns for travelers under CAV scenarios. Our model accounts

for a full spectrum of daily activity decisions and simulates these behaviors within a large transportation network to assess system-wide impacts.

- This study enhances existing ABM modeling approaches on CAVs by incorporating the CAV road capacity enhancement to the SCAG traffic assignment model and integrating the assignment model with the CAV-incorporated ABM through a multi-loop equilibrium assignment process. CAV impacts on transportation system in Southern California, including VMT, emissions, and transportation equity are evaluated.
- This research designs a set of demand management strategies and policies based on SCAG’s travel demand management (TDM) strategy toolbox, aiming at resolving the issues caused by CAVs, and then assesses the effectiveness of reducing excessive VMT and emissions and resolving equity concerns.
- This study proposes planning level policy recommendations for future system planning and management in the era of CAV. Provided guidance on strategies that are yet to be implemented in the real world or have difficulty in quantifying the performance but still contribute to eliminating the negative impacts caused by CAV adoption.

The rest of this chapter is organized as follows: We introduce studies that analyze the impact of CAVs on road capacity, system-level metrics such as VMT and emissions, and equity in Section 2.2. Section 2.3 briefly reviews the CAV-ABM of Southern California in the Part one paper [99] and discusses the methodology of integrating the CAV-ABM with the capacity-sensitive CAV traffic assignment model. The simulation scenarios and key performance metrics are also introduced in this section. Section 2.4 illustrates and analyzes the simulation results and insights, and Section 2.5 concludes the paper and discusses further directions.

## 2.2 Related Work

One critical issue for the CAV technology is that vehicle automation is still at its early stage. Therefore, the lack of empirical data about the use of CAVs and associated implications is inevitable. The deployment of CAVs could cause a significant impact on both the demand side and the supply side. We have reviewed the studies focusing on the demand side influence of CAVs in the part-I paper. In this section, we will mainly review the papers exploring the supply side impact of CAVs. Many studies have investigated the impacts of CAVs on roadway capacity through different approaches such as traffic simulation, road test, driving simulators, among others [169]. Ni et al. [189] estimated the improvement that connected vehicles could have on freeway capacity by incorporating the effect of connected vehicles on car-following behavior. The simulation results indicated that the capacity could be improved by as high as 50% if the connected vehicles are fully in use [189]. Olia et al. (2018) developed an analytical model to study the influence of AVs on highway capacity. The model proposed by the authors considered both the car-following behavior and lane-change behavior and was evaluated in the micro-simulation. Multiple market penetration rates of AVs were simulated. The results show that a 300% increase in lane capacity can be realized under the 100% market penetration scenario [201]. Adebisi et al. [3] developed a quantified approach to evaluate the influence of CAVs on freeway capacity and proposed adjustment factors for the capacity values recommended by Highway Capacity Manual (HCM). The authors evaluated a series of scenarios by combining different freeway configurations and multiple CAV applications. The microscopic simulation results indicated that the capacity can be increased by as much as 40% based on test scenarios and CAV market penetrations [4].

Much literature have investigated how the surging travel demand due to CAVs affects the VMT by trip-based, activity-based, or agent-based simulation models. Fagnant and Kockelman [73] estimated the transportation and environmental effect of shared AV operation based on an agent-based simulation model. Multiple simulation tests were conducted by

considering factors such as fleet size, modeling area, level of traffic congestion, and so on. The simulation results indicated that compared with conventional vehicle cases, the total VMT could increase by 11% [73]. Childress et al. [45] incorporated the behavior changes induced by AVs into an activity-based travel demand model developed for the Seattle region. The authors proposed a series of assumptions to reflect the potential impacts that AVs might have on travelers and on the transportation system. The simulation results showed that there could be a 4% to 5% increase in VMT and a 3% to 4% decrease in vehicle hours traveled (VHT) by assuming a 30% growth in road capacity [45]. Auld et al. [15] estimated the implications of CAV technologies to the system mobility in an activity-based model incorporating AV behavior changes in the simulation platform POLARIS. The results indicated that VMT was increased by 10% to 40%, depending on the assumptions on the value of time and CAV market penetration rate [15].

Emission/fuel consumption implications of CAVs are also analyzed by many other studies based on simulations or field tests. Wadud et al. [256] proposed a coherent energy decomposition framework to investigate the underlying implications that automated vehicles might have on energy consumption and greenhouse gas (GHG) emission. The authors developed multiple scenarios by combining different conditions of AV market penetration rates, expected reductions in travel time, levels of congestion mitigation, among others. Modeling results from this study indicated that partial deployment of AVs could potentially reduce the emission, whereas fully automated traffic flow could induce a negative effect in terms of emissions [256]. Brown et al. [34] estimated the consumption and emission effect of vehicle automation using the MARKET Allocation (MARKAL) energy system model. The simulation results demonstrated that the surging demand induced by AVs could increase fuel consumption significantly. In the meantime, fuel prices could go down due to the higher energy efficiency in AVs compared to conventional vehicles [34]. Ma et al. [160] developed an optimization algorithm dedicated to controlling and evaluating the energy consumption of CAVs. The authors developed the model based on the Relaxed Pontryagin's Minimum Principle (RPMP) and conducted field

experiments on rolling terrains. The test results implied that an around 20% decrease in energy consumption can be achieved compared with a conventional driving system [160].

There are very few studies on how the introduction of CAVs impacts travel equity in the context of a large-scale transportation network. Milakis et al. [168] analyzed the potential effects of AVs on travel accessibility and equity by applying the Q method to a conceptual model proposed by Geurs and van Wee [83]. The authors also concluded that AVs could cause negative impacts on social equity due to the varied affordability of this technology [168]. Winter et al. [268] developed an agent-based model to investigate the mobility change with the introduction of shared AVs. By testing the scenarios with different shared AV relocating plans, the results showed that the heuristics balancing of demand and supply provided the most equity benefits to the users [268]. Nahmias-Biran et al. [182] developed a modeling framework using the activity and agent-based approach in three cities' networks to estimate the implications of AVs on equity. The model covered both auto-dependent and transit-oriented network patterns. Simulation results showed that the AV benefits on equity could be achieved in a transit-oriented network when the population size is large. On the other hand, auto-dependent networks induce much less equity enhancement under the AV environment [182].

Very few papers specifically addressed the CAV deployment impact problem using the combination of activity-based demand and traffic assignment model. Unlike existing studies [45,91], which often use assumptions to model CAV behavior shifts, our research employs both activity-based demand and traffic assignment models while incorporating stated preference data. Moreover, it addresses equity concerns explicitly, integrating demand-supply interactions to offer a comprehensive analysis of CAV impacts on system performance and equity.



## 2.3 Methodology

### 2.3.1 Integration of CAV-ABM and Assignment Model

#### CAV-ABM

This section reviews the CAV-ABM from He et al. [99], and readers can refer to the Part-I paper for the explicit description of model development and demand analysis findings. The CAV-ABM was developed based on the state-of-the-art ABM framework from SCAG [195]. The SCAG ABM models people’s travel behaviors in Southern California from long-term (e.g., work arrangement and location choices) to short term (e.g., mode choices), as shown in Fig 2.1. There are seven layers of this model. The first layer generates a synthetic population for the SCAG area with household-level demographic and socio-economic attributes. In the following layers, activity and tours would be generated and formulated according to synthetic person’s type (e.g., worker vs. non-worker). With the synthetic population as input, the second layer predicts people’s long-term choices, like work arrangements, usual school and work locations, work schedule flexibility, among others. After the long-term choices are made, the model determines the mobility choices in the third layer, i.e., driver license and vehicle ownership. In the fourth layer, people’s medium-term choices are modeled with the sequence of mandatory activities (work and school) and non-mandatory activities (household maintenance and individual discretionary). Time window constraints of activities are also considered. If a household has school-age child/children, the school escort activity is generated before non-mandatory activities. The fifth layer simulates the interactions among household members. In layer six, with all activities generated, tours are formulated to connect all the activities, and the departure time and mode choices are also determined afterward. Thus, the activity-based travel demand is generated. Lastly, the SCAG ABM assigns the activity-based travel demand to the physical network (road and transit) to estimate VMT and other performance metrics to support their transportation planning decision-making.

A stated-preference survey was designed and distributed in the six counties in Southern California to collect people’s willingness to use CAV and the travel behavior changes associated with the CAV deployment. Based on the survey data, a series of models (marked by red boxes in Fig 2.1) were re-calibrated to incorporate such CAV-associated choice behavior changes of CAV users in the SCAG ABM framework. In the long-term choice layer, the choices of work arrangement and usual work location were re-calibrated to incorporate the behavior changes of CAV users. In the mobility choice layer, a CAV ownership model was developed to predict the number of CAVs owned per household and presented as one of inputs to subsequent model layers. In the fourth layer, the adult mandatory activity frequency model was updated to capture people’s new choices with CAVs. Similar updates were made to maintenance and individual discretionary activity generation models. At last, the mode choices model was updated to incorporate both PAV and SAV as emerging mode choices.

According to the survey results, about 53% of the residents in Southern California would like to use CAVs for their daily travel, and CAV-users’ travel behaviors changed significantly regarding work locations, activity frequencies, and mode choices. The CAV-ABM also predicted that the total trip number would increase by 7% with an 11% growth in the travel distance by car-like modes.

### **Network Capacity Adjustment**

To incorporate the influence of CAV on roadway capacity in our network, we adopted the highway capacity adjustment factors from the study by Adebisi et al. [3,4]. The authors used the roadway capacity estimation developed by the Highway Capacity Manual (HCM) as the base model, then developed capacity adjustment factors (CAF) for CAVs. The authors considered different scenarios that could impact the capacity adjustment, including freeway configurations, traffic demand, and CAV market penetration rate (MPR). The impacts of CAVs on roadway capacity under different scenarios were evaluated to adjust the use of HCM. The study considered different freeway segments including basic freeway, merge, and freeway

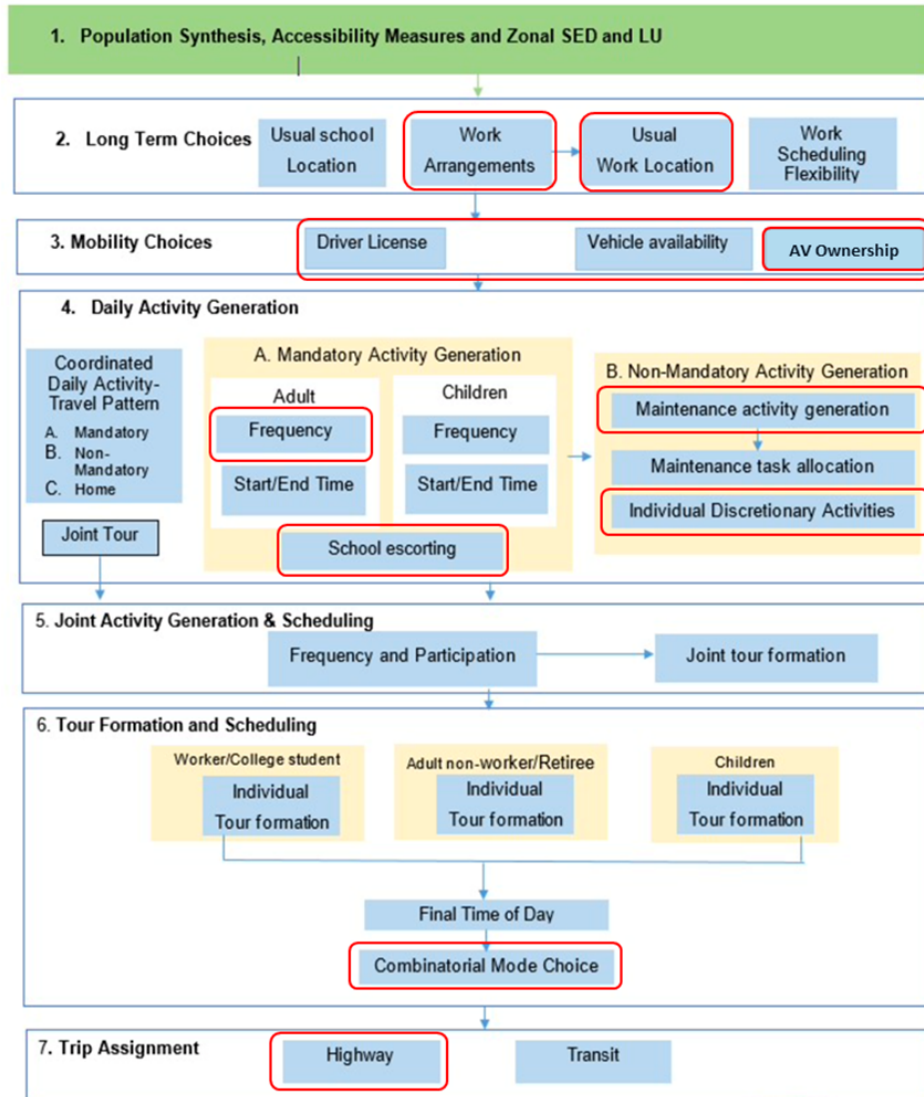


Figure 2.1: System design of SCAG ABM

weaving segments. Three traffic demand, or starting capacity were evaluated: 2,400, 2,100, and 1,800 veh/hr/ln. Five CAV MPRs was considered: 0%, 20%, 40%, 60%, 80%, and 100%. By conducting microscopic simulation in VISSIM, the study explored the enhancement on roadway capacity due to CAV under different scenarios and provides guidance to decision makers on the final capacities. In this study, we apply adjustment factors to all freeway and arterial links according to the HCM study. In the meantime, we use the proportion of CAV-mode trip distance in total car-mode trip distance to approximately represents the

MPR as Equation 2.1.

$$MPR = \frac{D_{CAV}}{\sum_{i \in M} D_i} \quad (2.1)$$

where  $M = \{CAV, SOV, HOV\ 2\ driver, HOV\ 3\ driver, taxi\}$ ,  $D_i$  refers to the total travel distance of each of the modes in  $M$ , and  $D_{cav}$  refers to the total travel distance of CAVs. Here, SOV means single-occupancy vehicles and HOV represents high-occupancy vehicles. Specifically, HOV 2 refers to HOVs with 2 occupants and HOV 3 refers to HOVs with 3 or more occupants. In Adebisi’s paper, the capacity adjustment factors under each CAV MPR are further differentiated into three values based on the link’s original capacity. This assumption considers the influence of the original link capacity on the link capacity adjustment factor. In this study, we only referred to the highest factors under each CAV MPR case so that the original capacity of each link is not considered when applying the capacity adjustment. In other words, we assumed the best-case scenario for the capacity improvement on network links under the CAV environment. According to the results in the Part-I paper, the proportion of trip distance of CAV-mode in all car-mode trip distance is 19.2%. Therefore, according to the capacity adjustment table developed by Adebisi et al. [3, 4], we selected 1.15 as the capacity adjustment factor in our model network.

### **Multiclass User Equilibrium Highway Assignment**

The multi-class highway assignment procedure adopted in this model simultaneously loads the trips predicted by the activity-based demand model and the three classes of heavy-duty trucks (HDT) denoted as light HDT (LHDT), medium HDT (MHDT), and heavy HDT (HHDT). The vehicle classes to be assigned to the network are as follows: Drive Alone, Shared Ride with 2 occupants, Shared Ride with 3 or more occupants, LHDT, MHDT, and HHDT.

To transform CAV mode trips generated from the demand model to the OD trip tables, all the PAV trips are aggregated into Drive Alone class and all the taxi and SAV trips

are aggregated with the Shared Ride 2 class. In the rest of this section, we will discuss the procedure of the highway assignment. Highway assignment represents the process of loading vehicles onto appropriate highway links. The outputs of the assignment include traffic volumes, congested speeds, VMT, and vehicle-hours traveled (VHT) estimates. In our model, the assignment is made for five different time periods, i.e., AM Peak (6 am to 9 am), Midday (9 am to 3 pm), PM Peak (3 pm to 7 pm), Evening (7 pm to 9 pm), and Night (9 pm to 6 am). The assignments on links or segments of different time periods are integrated to generate the average daily traffic volumes (ADT) for the model network. The model road network is shown in Fig 2.2.

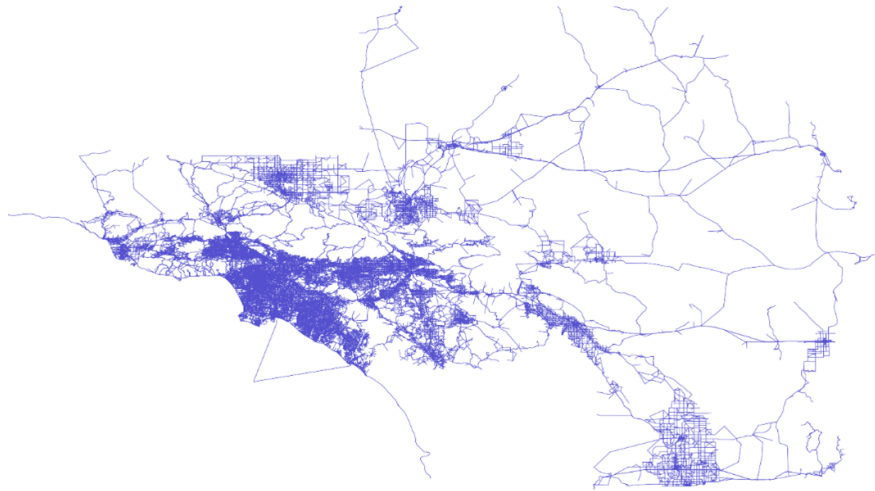


Figure 2.2: Model road network in Southern California

The traffic assignment module in the SCAG 2016 model includes a set of multi-class simultaneous equilibrium assignments for the eight abovementioned classes of vehicles at each time period. The model involves a three-loop feedback process so that the convergence of travel time can be achieved between the demand model and the highway assignment model [195]. There are two reasons why we chose to use three loops: 1) in most cases the congested speed barely changes after three loops; 2) running the SCAG ABM is very time-consuming, and a single loop takes around 24 to 36 hours to finish using the work

station for this study. In the meantime, 10% population sampling was used in the demand model to save computation time when testing multiple scenarios. 10 percent of the total households are sampled based on household ID, then household members corresponding to the sampled households are selected as the synthetic population input for the demand model. The total computation time is reduced to three days compared with six to seven days with the full population sampling condition. Though the demand is generated based on 10% population, the OD trips are scaled up to reflect the demands under full population. Therefore, the assignment results still reflect the impact of the full population demand. Fig 2.3 illustrates how the demand model is combined with the assignment model and how the feedback between each model loop interacts. The process is described as follows:

**Step 1:** When starting the demand forecasting during the first loop, the travel times are derived from the speeds coded on the input highway networks. The demand model then generates the OD trip tables that contain trips of each vehicle class at different time periods. The trip tables created for different vehicle classes at each time period are then assigned to the highway links. This process creates the first set of traffic volumes and congested speeds.

**Step 2:** The congested speeds generated from the assignment process in the previous loop are further used as the inputs for the demand model in the following loop. The updated demand will result in new OD trip tables and then be loaded to the network links to create a new set of congestion speeds. The variation of volume between the two assignment loops is smoothed by an approach widely used in traffic assignment called successive averages. The step size used for the averaging method is  $1/i$ , with  $i$  being the number of iterations, and the equilibrium criteria is when the relative gap is less than 0.001 or the averaging process has reached 200 iterations.

**Step 3:** The travel times processed by the successive averages in the previous loops are further input into the final loop, where the last assignment is conducted for all time periods.

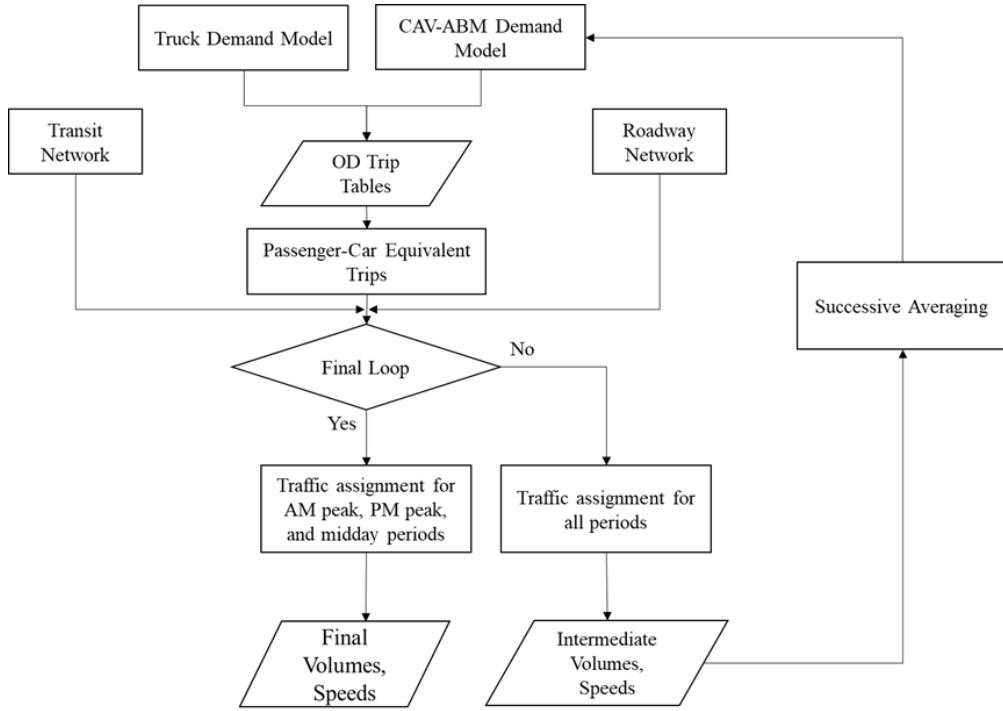


Figure 2.3: Model Integration Framework

### 2.3.2 Scenario Design

In the part-I paper of this study, we already analyzed the demand side impacts by introducing CAVs into the model [99]. With the deployment of CAVs, we can expect around a 7% increase in total trip numbers and about an 11% increase in total travel distance for car-like mode trips. With that being said, the influence of the induced demand on the transportation network incorporating the supply side consideration has yet been thoroughly investigated. The potential downside impacts, such as increased VMT, increased emission, congestion, and travel equity concerns, also need to be further discussed based on the results from different test scenarios.

#### CAV Base Case Scenario

The base case scenario is developed based on the CAV-incorporated travel demand model, the capacity adjusted network described, and the traffic assignment procedure presented in

Section 2.3.1. By assigning the demand of both CAV and non-CAV users into the network, the base case scenario generated the baseline outputs that reflect the system-level impacts. The outputs of the base case scenario will be compared with the original SCAG model output, so that we can quantitatively evaluate the impacts of the CAV deployment. Note that no travel demand management policies were applied in the base case scenario.

## **Demand Management Policy Scenarios**

Building on the base case scenario outlined in Section 2.3.2 and the travel demand management (TDM) toolbox detailed in the SCAG TDM strategic plan report [194], we developed a set of TDM policies aimed at mitigating the negative impacts of CAV integration on the transportation system. In the SCAG TDM report [194], the TDM strategies that already exist or will potentially be applied in the SCAG area can be grouped into the following categories:

- Parking strategies include multiple parking pricing policies aiming to increase the cost of single-occupancy driving.
- System enhancement strategies produce competitive modes as alternatives to driving such as transit, taxi, shared bikes, etc. The improvements include both physical (e.g., transit lines extension) and systems (e.g., lower transit fares) upgrades.
- Incentives and Facilitation strategies include providing monetary benefits to travelers for using alternative modes or offering flexible work schedules or telework options to reduce commuting trips.
- Education and Marketing strategies involve providing travelers with educational information regarding travel options or launching marketing programs to promote people's awareness of traveling with alternative modes.
- Other supportive strategies.



This study formulated test policies for each of these categories, taking into account the feasibility of implementing the scenarios within our model. For example, since the model cannot simulate infrastructure upgrades to the transit network, we focused on transit improvements from a pricing perspective, such as offering subsidized transit passes. To assess the effectiveness of these policies, we designed multiple scenarios, incorporating the policy elements into the base CAV demand model and evaluating them through the integrated demand-assignment framework.

*Scenario 1: Telework*

Telework and remote travel are direct and effective ways to reduce SOV trips and VMT. This scenario is inspired by SCAG’s "Future of the Workplace" program [194], which analyzed the characteristics of telecommuting behavior of those who work from home and those who travel to remote worksites. In the real world, this requires employers to provide opportunities for employees to work remotely from the workplace, thereby helping to eliminate the need for them to travel to the office or worksite. In this scenario, we adopted an optimistic assumption that the ratio of workers who work from home is doubled across households from different income groups. This change was made in the long-term decision module in the demand model where we doubled the target work from home rate for each of the seven income groups as shown in Table 2.1.

<b>Annual Household Income</b>	<b>SCAG Model</b>	<b>Scenario 3</b>
< \$25,000	8.9%	17.8%
\$25,000 to \$50,000	6.4%	12.8%
\$50,000 to \$75,000	6.3%	12.6%
\$75,000 to \$100,000	6.2%	12.4%
\$100,000 to \$125,000	6.7%	13.4%
\$125,000 to \$150,000	8.5%	17.0%
\$150,000 to \$200,000	9.4%	18.8%
> \$200,000	10.6%	21.2%

Table 2.1: Work from Home Rate by Income Groups in SCAG Model

*Scenario 2: Transit Fare Subsidy*

In this scenario, we encourage travelers to use more transit modes as alternatives to the driving mode by providing subsidies on transit passes. Similar policies have already been put into practice in different regions around the U.S. Some cities are seeking to provide subsidies according to income levels, such as the "LifeLine Pass" in San Francisco [237], thereby addressing both equity goals and TDM goals. To model this scenario, we assumed the best-case scenario where all the population is provided with free transit passes and the commute to the transit stations was subsidized. This is realized by setting the transit fare as well as the fuel cost of two car-related transit access modes, kiss-and-ride (KNR) and park-and-ride (PNR) to zero in the mode choice model.

*Scenario 3: Parking Pricing*

The first scenario assumes that parking costs for SOVs and PAVs are doubled, while the parking and the toll fees for HOVs are reduced by half. This scenario is designed to encourage the use of shared ride modes either HOVs or SAVs and other modes. To test the scenario, we changed the parking fare in the mode choice model [195]. In the base model, the utility of parking fare was defined as follows:

$$U_{\text{park}} = \text{cost scale} \cdot \beta_{\text{park}} \cdot \text{parking cost} \quad (2.2)$$

$$\text{cost scale} = \frac{1}{(\text{income}/1000)^{0.8} \cdot \text{party size}^{0.6}} \quad (2.3)$$

where the party size of the trip, parking cost is the model default parking fare related to parking duration, and  $\beta_{\text{park}}$  is the calibrated coefficient. The cost scale in Equation 2.2 is a scaling factor corresponding to income and the party size of the trip. Equation 2.3 shows the definition of the cost scale factor, where income refers to annual household income and party size represents the number of household members taking part in the trip. The utility function and cost scale factor were defined in the SCAG mode choice model [255], and the parameters were calibrated based on 2010-2012 California Household Travel Survey data [192]. The cost

scale is defined as a function of household income and trip party size and can be regarded as an indicator of people’s sensitivity to the parking cost.

#### *Scenario 4: Auto trade-in*

The last scenario assumes that CAV ownership households use purchased PAVs to replace the same number of regular vehicles. This policy is inspired by the Clean Cars 4 All program by California Air Resources Board [27] and Tesla’s Trade-Ins service (Tesla, Inc., 2022). To model that scenario, household auto ownership was changed. In the original model, the household auto ownership was modeled before the CAV ownership, which means the output of CAV ownership would not affect the auto ownership. To incorporate the "trade-in" effect after purchasing CAVs, we designed an auto ownership adjustment model after the CAV ownership model to reflect the impact of the number of CAVs on a household’s existing regular vehicles. We assumed that regular vehicles were equally replaced by CAVs to reflect a best-case scenario for all CAV ownership households.

### **2.3.3 Performance Metrics**

#### **System Impact Metrics**

In this study, we analyzed the implications of CAVs on the transportation system by multiple performance indicators, including VMT, vehicle hours traveled (VHT), traffic speed, mode share, trip length, etc. These metrics are widely used in large-scale transportation analysis [15, 45, 195]. The total VMT presented in the results and discussion section is the adjusted total VMT based on the raw model output to reflect the carpool effect in the SAV trips. In the CAV-incorporated ABM developed in our study, all SAV trips were counted as single-occupied trips, which might result in overestimated VMT from SAV as ride-sharing trips were double-counted. Therefore, we adopted an occupancy adjustment factor of 1.4, as suggested by [105], to adjust the VMT generated by SAV trips to obtain the realistic value.

The traffic speed is measured by the mean speed during two peak hour periods: AM peak

and PM peak. We also investigate the spatial congestion distribution during peak hours by measuring the ratio of mean speed to free-flow speed (FFS) defined as follows:

$$R = \frac{V_{\text{mean}}}{V_{\text{FFS}}} \quad (2.4)$$

where  $V_{\text{mean}}$  is the average speed of the link during the selected time period;  $V_{\text{FFS}}$  is the free-flow speed of the corresponding link.

Emissions are calculated based on the emission factors proposed by EMFAC [48]. The emission factors are used to convert the traffic output, such as VMT, into equivalent emission output. The full version of emission factors depends comprehensively on a series of attributes like region, fuel type, vehicle categories, etc. The final emission output in our model is the sum of the emissions from 6 counties in the SCAG region. We select four major pollutants as the emission analyzing objects: Carbon Dioxide (CO<sub>2</sub>), Nitrogen Oxide (NO<sub>x</sub>), Particulate Matter 2.5 (PM<sub>2.5</sub>), and Reactive Organic Gases (ROG).

## Equity Metrics

Aside from system-level impacts, we performed equity analysis for each test scenario at the household level. Equity is a measure of the distribution of outputs (or inputs) across the population in a fair manner [143]. We analyzed the equity impacts from three aspects: (1) trip numbers per household, (2) travel distance per household, and (3) household travel accessibility. These equity metrics are adopted by another study regarding equity implications of AVs [182]. For each household, the travel accessibility is calculated as the log sum of the utilities across all modes for each tour and averaged across tours generated from this household, as shown in Equation 2.5.

$$A = \frac{\sum_{j=1}^n \log \sum_{i=1}^{m_j} e^{u_{ij}}}{n} \quad (2.5)$$

where  $n$  is the number of tours of the household,  $m_j$  is the number of available mode combinations for tour  $j$ , and  $u_{ij}$  refers to the sum of trip utilities of the  $i$ th available trip mode combination for tour  $j$ , which can be expressed as:

$$u_{ij} = \sum_{k=1}^{K_j} v_{kij} \quad (2.6)$$

where  $v_{kij}$  is the utility of trip  $k$  in mode combination  $i$  for tour  $j$ , and  $K_j$  is the number of trips in tour  $j$ .

Households are categorized into three groups by annual income: lower, middle, and upper. We use relative values instead of absolute values to evaluate the implications of equity metrics. In each analyzed scenario, outputs of the lower-income household are used as the baseline (denoted as 1.00), and outputs of middle- and upper-income households are represented in relative values proportional to the baseline.

An equity index is used to represent the disparity of equity measurements across income groups in a test scenario, defined as:

$$\text{Equity Index} = \frac{\text{Range}\{L, M, U\}}{\text{Range}\{L_0, M_0, U_0\}} \times \begin{cases} -1, & \text{if } \max\{L, M, U\} = L_i \\ 1, & \text{if } \max\{L, M, U\} \neq L_i \end{cases} \quad (2.7)$$

where  $L, M, U$  refer to the value of equity measurements of lower-, middle-, and upper-income households in tested CAV scenarios, while  $L_0, M_0, U_0$  denote the value of equity measurements in the SCAG model.

## 2.4 Results and Discussion

### 2.4.1 Transportation System Impacts Analysis

The simulation outputs of the SCAG model, the CAV base model, and Scenarios 1 through 4 are presented in Table 2.2. As shown in Table 2.2, VMT, number of trips, and trip distances have a significant increase in all the CAV models. This indicates that the induced demand from the deployment of CAVs generated more trips and therefore increased the miles traveled in the network. Compared to the base SCAG model, total VMT increased by 10.0% following the adoption of CAVs. This suggests that people are more inclined to travel longer distances with CAVs. This finding is consistent with results from existing studies on the impact of CAVs on travel behavior. For instance, Fagnant and Kockelman [73] estimated the transportation and environmental effects of shared AV operations using an agent-based simulation model. Their multiple simulation tests, which considered factors like fleet size, modeling area, and traffic congestion levels, showed a potential VMT increase of 11% compared to conventional vehicles. Similarly, Auld et al. [15] assessed the impact of CAV technologies on system mobility using an activity-based model within the POLARIS simulation platform. Depending on assumptions about the value of time and CAV market penetration rates, their results indicated a VMT increase ranging from 10% to 40%.

The average speed on freeways during AM peak hours has increased after adopting CAVs, while it barely changed at PM peak and on arterials. The number of trips during the PM peak increased significantly, which offset the improvement of road capacity and resulted in slight changes in travel speed. This finding is also consistent with the lower increase rate in VHT than in VMT, since vehicles have a higher speed on freeways during AM peak hours. As for the arterials, the traffic speed has only been increased slightly during the AM peak hours, whereas the PM speed on arterials is even decreased in the CAV model. This indicates that the induced demand on arterials has offset the capacity enhancement due to CAVs, which caused worse congestion during the PM peak hours for arterials.

Metrics	Value	SCAG	CAV Base	CAV S1	CAV S2	CAV S3	CAV S4
<b>VMT (mi)</b>	Daily VMT	$4.57 \times 10^8$	$5.03 \times 10^8$	$4.93 \times 10^8$	$4.98 \times 10^8$	$4.97 \times 10^8$	$4.71 \times 10^8$
<b>VHT (h)</b>	Daily VHT	$1.39 \times 10^7$	$1.47 \times 10^7$	$1.42 \times 10^7$	$1.44 \times 10^7$	$1.43 \times 10^7$	$1.31 \times 10^7$
<b>Trips</b>	Number of trips	12.8	13.84	13.82	13.85	13.85	13.31
<b>Trip length (mi)</b>	Mean trip length	7.18	7.17	7.10	7.17	7.18	7.21
<b>Speed (mph)</b>	Freeway AM speed	52.79	55.88	56.24	55.91	55.86	56.77
	Freeway PM speed	52.82	53.09	53.43	53.29	53.22	54.27
	Arterial AM speed	29.76	29.84	29.95	29.90	29.93	30.21
	Arterial PM speed	29.52	29.11	29.26	29.19	29.21	29.61
<b>Daily Emission (ton)</b>	NOx	26.57	29.20	28.65	28.95	28.82	27.34
	PM2.5	0.67	0.74	0.73	0.73	0.73	0.69
	CO2	$1.39 \times 10^5$	$1.53 \times 10^5$	$1.50 \times 10^5$	$1.52 \times 10^5$	$1.51 \times 10^5$	$1.43 \times 10^5$
	ROG	6.61	7.25	7.11	7.18	7.15	6.79

Table 2.2: Comparison of Metrics Across Different Scenarios

## 2.4.2 Results and Discussion

The spatial distribution of roadway speed is illustrated in Fig 2.4 and 2.5. In this study, we used the ratio of mean speed to free-flow speed to demonstrate the congestion status during different time periods. Since most of the congestion occurred in LA County, we presented the speed distribution of LA County exclusively. Note that in both Fig 2.4 and 2.5, (a) and (c) are the general views of LA County, while (b) and (d) are the zoom-in views of the downtown LA area. Fig 2.4 displays the comparison of AM congestion status between the SCAG model and the CAV base model. As shown in blue circles, the congestion mainly occurred on freeways and arterials near the downtown area, such as I-405, I-10, and US-101. With the deployment of CAVs, the congestions are well reduced on multiple freeways around downtown LA, as marked by the dashed circles in Fig 2.4b and Fig 2.4d. However, the traffic speed on arterials does not significantly increase after using CAVs. This trend matches the arterial speed results shown in Table 2.2, indicating that the induced demand on arterials has offset the capacity enhancement for arterials. Conversely, Fig 2.5 provides a view of the PM traffic speed comparison between the SCAG and the CAV model. Compared with the

results in Fig 2.4, the PM period shows worse traffic congestion than in AM. Furthermore, the enhancement of traffic speed with CAV deployment during the PM period is not as significant as in AM, as fewer freeways and arterials show congestion reduction with the use of CAVs in Fig 2.5.

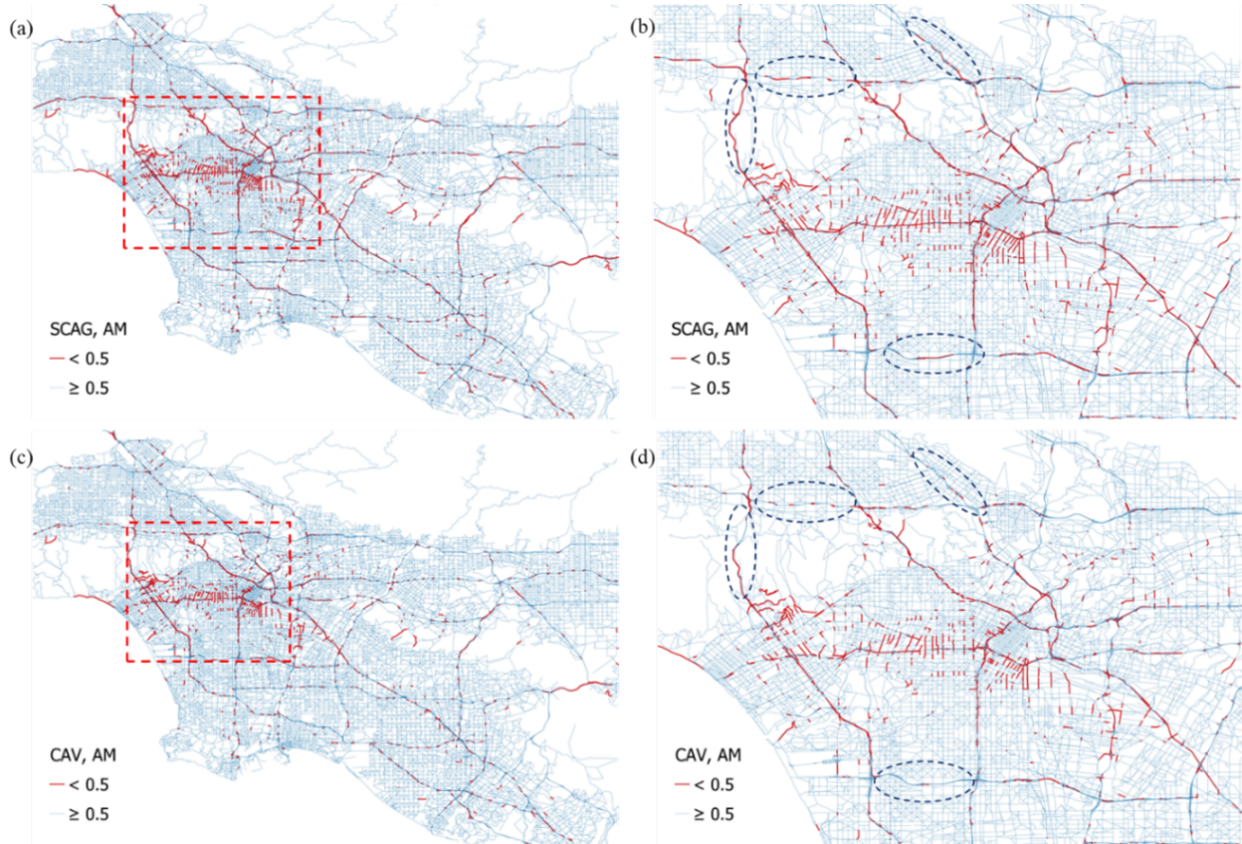


Figure 2.4: Ratio of mean speed to free-flow speed during AM peak: (a) SCAG; (b) Zoom in view of (a); (c) CAV; and (d) Zoom in view of (c).

The results of mode share change before and after CAV deployment are depicted in Fig 2.6. The CAV demand-only refers to the mode share outputs directly from the demand model in part-I paper (He et al., 2022), which did not incorporate the 3-loop assignment feedback introduced in this study. The CAV demand and supply refers to the final outputs after considering the feedback effect of the assignment results after 3 loops of iteration. As indicated from Table 2.2 and Fig 2.6, the mode share shows a clear shift from non-CAV modes



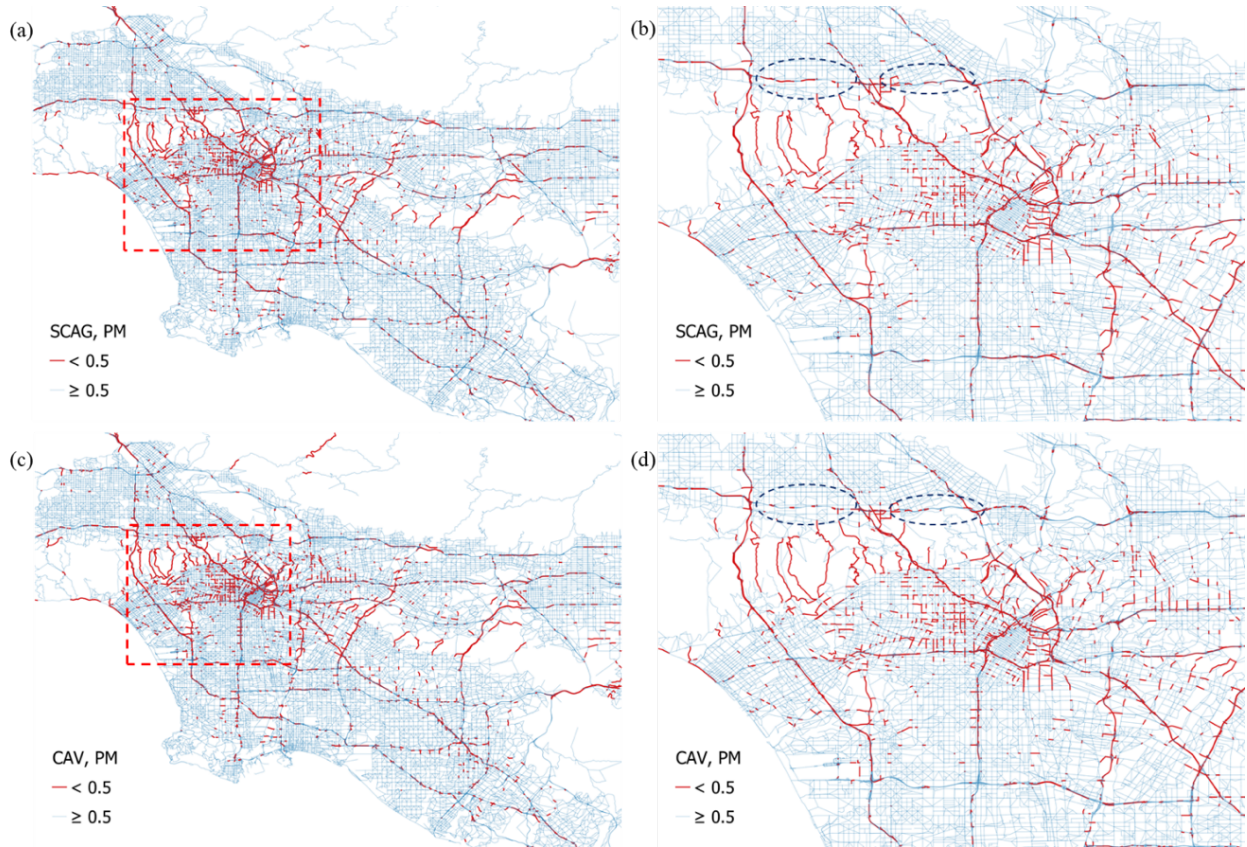


Figure 2.5: Ratio of mean speed to free-flow speed during PM peak: (a) SCAG model; (b) Zoom in view of (a); (c) CAV base model; and (d) Zoom in view of (c).

to CAV modes and transit before and after CAV deployment. Fig 2.6 also reveals that the outputs of demand-only slightly overestimated the share of SOV and HOV. It is noteworthy that the mode shift from SOV to CAV is smaller than from HOV to CAV, suggesting that most SOV drivers tend to maintain their travel habits while HOV drivers and passengers are more willing to switch to CAV modes. Note that the transit mode share has increased from 1.97% in the SCAG model to 6.3% in the CAV model. This growth trend in transit mode share aligns with [111], where the transit mode share had a 10 times increase after introducing AVs. Although we did not model the first- and last-mile trips to transit stations using CAVs, i.e., kiss-and-ride (KNR) and park-and-ride (PNR), this part of the growth in transit share can be explained as alternative trips by users who intended to ride CAVs but

did not have available ones at the moment before their trips take place. Considering the small size of transit trips, the mode share predicted from the model might also be influenced by the stochastic noise of the simulation.

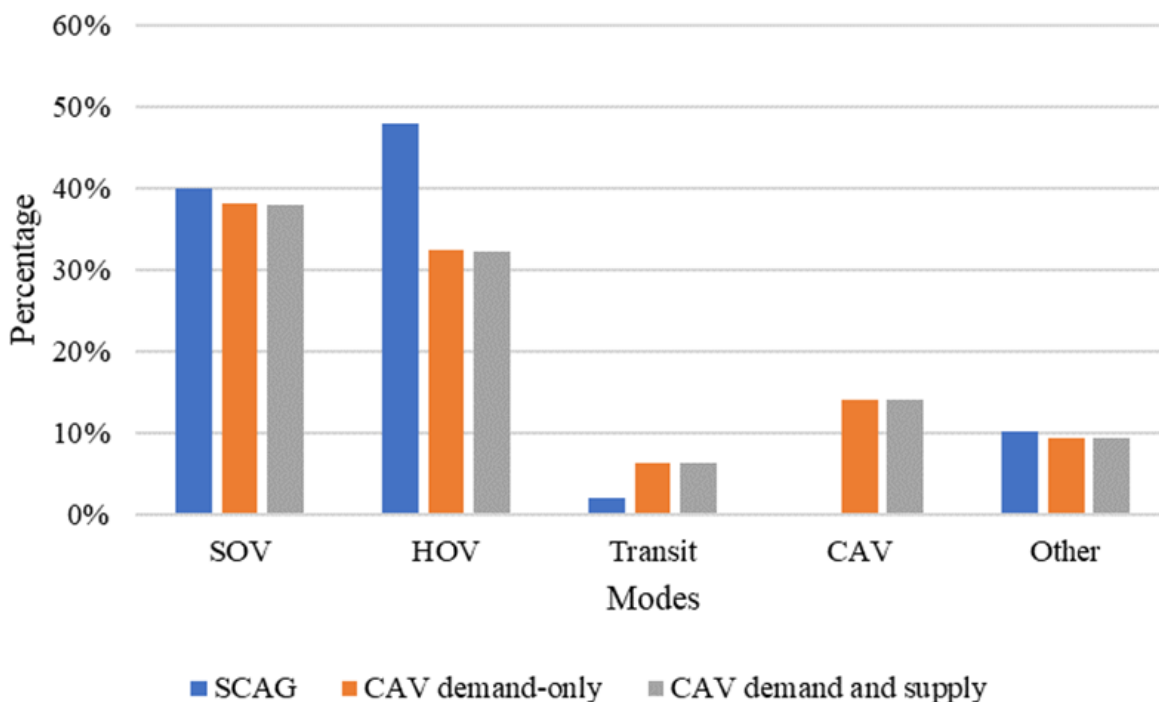


Figure 2.6: Comparison of mode share between SCAG, CAV demand-only, and CAV demand and supply.

The emission results from Table 2.2 indicated very similar trends as the VMT results. Compared with the SCAG scenario, all the CAV scenarios have shown an increasing trend in emission for all four pollutant types. The CAV base scenario causes 9% to 10% of emission growth compared with the SCAG model. After applying the management policies, the CAV models indicate a reduction in emissions, ranging from 1% to 6%. The policy that induced the most emission drop is the auto trade-in from Scenario 4. The telework policy and parking pricing management policies caused less decrease in emissions, while the free transit policy had the least impact on reducing emissions.

### 2.4.3 Travel Equity Analysis

The equity performance measured by relative trip numbers and relative travel distance per household are presented in Table 2.3 and 2.4. Meanwhile, the comparison of equity indices for these two equity metrics is shown in Fig 2.7. As seen from Table 2.3 and 2.4, the disparity in trip numbers and travel distance across different income groups exists both in the SCAG model scenario and in all the CAV scenarios: lower-income households have fewer trips under all purposes than medium and upper-income households. As shown in Fig 2.7, all the CAV scenarios have lower equity indices in mandatory trips but higher equity indices in non-mandatory trips compared with the SCAG scenario. This indicates that the use of CAVs helps to reduce the equity issues when traveling for mandatory activities, whereas it worsens the disparity in household travel demand generation when it comes to non-mandatory trips.

Model	Mandatory				Non-Mandatory			
	Lower	Middle	Upper	Equity Index	Lower	Middle	Upper	Equity Index
SCAG	1.00	1.43	1.51	1.00	1.00	1.29	1.20	1.00
CAV Base	1.00	1.39	1.46	0.92	1.00	1.35	1.37	1.28
CAV Scen 1	1.00	1.39	1.45	0.88	1.00	1.35	1.37	1.26
CAV Scen 2	1.00	1.39	1.46	0.91	1.00	1.35	1.37	1.28
CAV Scen 3	1.00	1.40	1.47	0.92	1.00	1.35	1.37	1.28
CAV Scen 4	1.00	1.42	1.49	0.97	1.00	1.34	1.37	1.28

Table 2.3: Equity Performance Measured by Relative Number of Trips per Household

Scenario	Mandatory				Non-Mandatory			
	Lower	Middle	Upper	Equity Index	Lower	Middle	Upper	Equity Index
SCAG	1.00	1.78	2.07	1.00	1.00	1.32	1.29	1.00
CAV Base	1.00	1.74	1.96	0.90	1.00	1.40	1.49	1.54
CAV Scen 1	1.00	1.73	1.91	0.85	1.00	1.39	1.48	1.50
CAV Scen 2	1.00	1.75	1.97	0.90	1.00	1.40	1.49	1.55
CAV Scen 3	1.00	1.75	1.97	0.90	1.00	1.40	1.49	1.54
CAV Scen 4	1.00	1.75	1.97	0.90	1.00	1.39	1.48	1.51
CAV Scen 3 *	1.00	1.75	1.97	0.90	1.00	1.40	1.49	1.54
CAV Scen 4 *	1.00	1.39	1.48	0.44	1.00	1.37	1.44	1.38

Table 2.4: Equity Performance Measured by Relative Travel Distance per Household

Comparing Scenario 1 through Scenario 4 with the SCAG scenario and CAV base scenario in Table 2.3 and 2.4, and Fig 2.7, we can see that Scenario 1 has the lowest equity index among

Scenario	Income Group			Equity Index
	Lower	Middle	Upper	
SCAG	1.00	1.03	1.05	1.00
CAV Base	1.00	1.03	1.04	0.93
Scen 1	1.00	1.02	1.04	0.84
Scen 2	1.00	1.03	1.04	0.90
Scen 3	1.00	1.03	1.05	1.04
Scen 4	1.00	1.01	1.02	0.52
Scen 3 *	1.00	0.98	0.98	-0.52
Scen 4 *	1.00	1.02	1.02	0.48

Table 2.5: Equity Performance Measured by Relative Travel Accessibility

all scenarios both in trip numbers and travel distance, indicating the telework policy provides more equity improvement in terms of household trip numbers and travel distance than others. This is mainly because the reduction of commute trips is more significant in middle- and upper-income groups since the work-from-home rate is higher in these two groups, i.e., the demand increase in high-income households is less than that in lower-income households. The equity index of Scenario 2 in both relative trip numbers and relative travel distance are close to that of the CAV base model, indicating that the free transit pass policy does not cause significant implications from the equity perspective. In other words, the population is not sensitive to the cost change in transit across income groups, indicating that monetary policies cause a limited impact on changing people’s travel patterns, given the assumption that the transit network remains unchanged. Scenario 3 shows a very close equity index to the CAV base scenario, which indicates that differentiating the parking price for SOV and HOV does not influence the disparity in terms of demand generation and travel distance. The equity index of Scenario 4 in trip numbers per household is the highest among all test scenarios, very close to the value of the SCAG scenario, where no CAVs are considered. This indicates

that the auto trade-in policy offset the equity improvement in the CAV base scenario. This can be explained by the fact that although a generic auto trade-in reduces the total auto ownership, i.e., both AVs and non-AVs for all income groups, higher-income households can still meet the increased demand due to more AVs purchased, whereas lower-income groups might experience a lower demand increase because of the fewer AVs they owned. With that being said, a specified auto trade-in policy in favor of lower-income households instead of a generic auto trade-in policy should be tested out to ensure that the equity performance is not sacrificed.

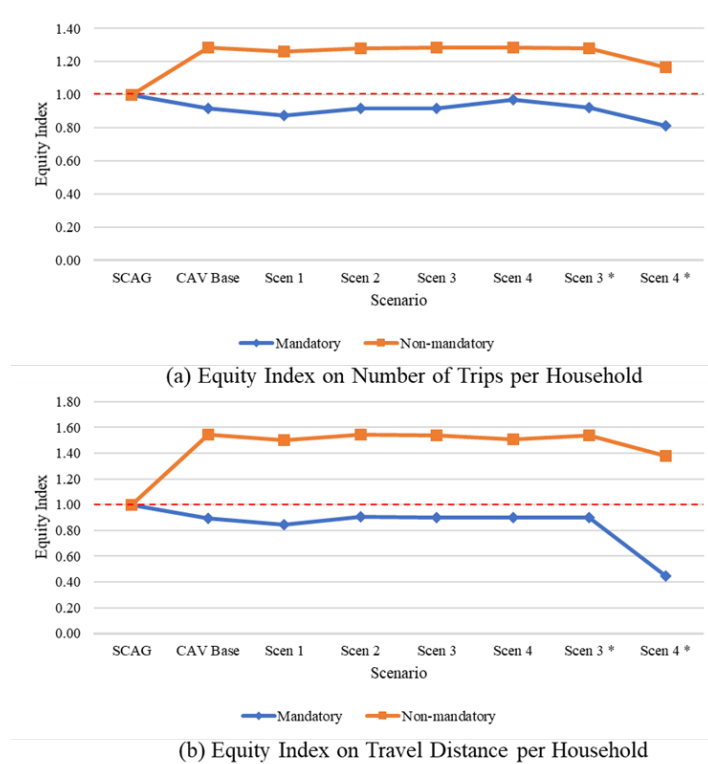


Figure 2.7: Equity Index across Scenarios for (a) Number of Trips per Household and (b) Travel Distance per Household

The equity performance measured by household travel accessibility and corresponding equity indices for each test scenario are presented in Table 2.5 and Fig 2.8. Table 2.5 illustrates that the equity index of CAV base, Scenario 1, 2, and 4 are smaller than the SCAG scenario, indicating a reduction in the disparity of travel accessibility across income

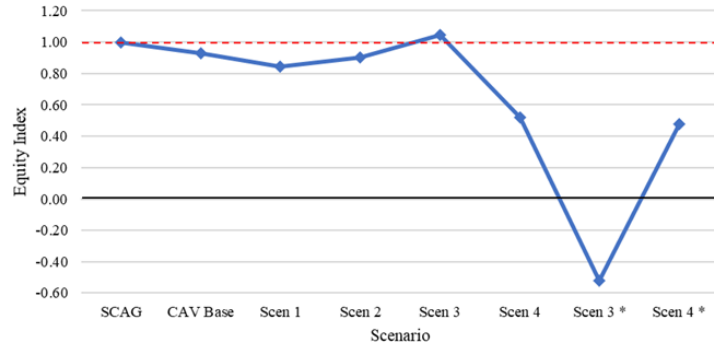


Figure 2.8: Equity Index across Scenarios on Travel Accessibility

groups in these scenarios. However, it is notable that the equity index for Scenario 3 exceeds one, suggesting that the equity performance in terms of travel accessibility deteriorates after implementing the parking pricing policy under CAV deployment. The main reason for this decline in equity is the variable sensitivity to parking costs across different income groups. As previously mentioned, an individual’s sensitivity to parking costs is modeled by the cost scale, which depends on household income; higher income results in lower sensitivity to price changes. Consequently, increased charges for SOVs and PAVs diminish lower-income households’ willingness to choose drive-alone modes, while not significantly deterring middle and upper-income households from using single-occupancy modes.

As concluded in Section 2.4.2, most CAV test scenarios show enhanced equity performance in terms of household trip numbers, household travel distance, and household travel accessibility compared with the original SCAG scenario. However, the equity of travel accessibility in the parking pricing scenario and the equity of household trip numbers in the auto trade-in scenario do not perform as well as in the CAV base scenario, indicating the need for further policy adjustments. Subsequent sections propose and evaluate modified policies for these scenarios by applying differentiated strategies across different income groups.

## 2.5 Conclusion

Recent studies have been investigating the potential impacts that CAV technologies could have on future travel. The deployment of CAVs will significantly change the way people live, travel, and participate in various activities. Therefore, understanding the implications of CAV deployment is important for stakeholders to prepare for future changes. In the part-I paper [99] of this study, we investigated the impacts of CAVs on people’s long-term to short-term travel behaviors by conducting a stated preference survey in Southern California and integrating the survey results with the SCAG ABM. The new travel demand with the deployment of CAV was estimated. This study went one step further based on the demand model generated from the part one study. We incorporated the supply-side impact of CAVs into traffic assignment modeling, integrated the travel demand model with the SCAG highway assignment model, and generated system-level outputs by considering the demand-supply interaction in feedback loops. We also proposed and evaluated a set of travel demand management (TDM) policies in response to the impacts of CAVs on the transportation system and analyzed the equity concerns that could occur with CAV deployment. Major findings and conclusions from this study are summarized as follows:

- The VMT and emissions increased significantly with the deployment of CAV. Given the base year model, 2016 SCAG ABM, and a model calibration by survey-collected CAV users’ data (53% of the population are CAV users), the predicted travel demand could lead to a 9.1% increase in VMT and a 9.6% to 10.4% increase in emissions.
- Variable parking pricing, telework promotion, and auto trade-in are effective strategies in reducing excessive growth of VMT and emissions induced by CAV adoption. On the other hand, VMT and emissions are not sensitive to transit fare reductions, even if transit is free of charge.
- Auto trade-in policy is the most effective policy in controlling the VMT growth in

future CAV deployment. Under the ideal condition where all CAV users follow trade-in rules, the VMT growth rate can be reduced from around 9% to 2%.

- With the deployment of CAVs, travel equity in terms of household trip number, household travel distance, and travel accessibility across income groups is promoted. The implications of the proposed TDM policies on travel equity are different. Telework helps to improve equity in all three metrics; free transit policy does not influence the equity performance; parking pricing policy tends to reduce the equity of travel accessibility, although it only has a limited impact on the equity of trip numbers and travel distance; auto trade-in policy worsens the equity in household trip numbers but helps promote equity in travel accessibility.
- Adjusted demand management policies based on income groups perform better in promoting transportation equity than generic policies. After equity adjustment, the trip number disparity in the auto trade-in scenario and the accessibility disparity in the parking pricing scenario are both reduced. However, excessive policy support for lower-income households can cause an overbalance in equity, which might need to be avoided.

There are areas to improve upon the work in this study. First, capacity improvement was a deterministic parameter and identical for all freeways and arterials in this study. In future studies, the capacity adjustment factor can be included in the loop feedback and evolved with the assignment results. Second, the proposed models in this study did not explicitly consider empty trips of CAVs, such as finding parking locations or returning home. This might become an inevitable question when the CAV market penetration rate gets higher in future CAV deployment. Third, the model can be further updated to simulate first- and last-mile service by CAV in Southern California and provide more explicit policy recommendations for public agencies to improve the existing transit system.



## Chapter 3

# Evaluating the traffic and environmental impact of telework and remote services on transportation systems

The proliferation of Information and Communications Technologies has democratized telework and teleservices (Tele-X), accelerated by evolving attitudes following the COVID-19 pandemic. As post-pandemic society increasingly embraces remote activities, Tele-X's potential to mitigate congestion and emissions remains uncertain. This study investigates the transformative impacts of Tele-X on travel patterns and assesses its potential to reduce Vehicle-Miles-Traveled (VMT) and traffic-related emissions. Our integrated modeling framework encompasses survey data, travel demand and supply modeling, and environmental analysis. We collected data regarding individual Tele-X choices via an online survey in Los Angeles (LA) County, which facilitates calibration of an activity-based model. The calibrated model reveals a rise in telework adoption from 6% to 24%. Despite marginal increases in total trips, post-pandemic VMT decreased by 3.9%, concomitantly affecting emissions. This underscores Tele-X's potential in reducing VMT and emissions. However, a case study within LA's Westwood community reveals that extreme promotion of Tele-X would induce passing-by traffic and

increase emissions in neighboring areas.

## 3.1 Introduction

The ongoing evolution of Information and Communication Technology (ICT) has instigated remarkable transformations within the digital world, thereby fostering a notable shift towards “everything” online. In alignment with the paradigmatic transition from the physical to the virtual realm, we introduce a general concept “Tele-X.” Tele-X encompasses a diverse array of activities facilitated by ICT, including telework and various teleservices. The term “telework” refers to “a work flexibility arrangement under which an employee performs the duties and responsibilities of such employee’s position, and other authorized activities, from an approved worksite other than the location from which the employee would otherwise work,” defined by the Telework Enhancement Act of 2010 [53]. Teleservice, or remote service, entails “the delivery of services from a distance using telephony and/or digital technologies” [52], such as online shopping, online education, telemedicine, online banking, etc. The transition to Tele-X engendered a reduction in physical travel (either commuting or non-commuting), providing a promising solution to alleviate adverse impacts on transportation and environment, such as traffic congestion, energy use, and traffic-related emissions [11, 54, 290].

Tele-X has become increasingly prevalent in people’s daily lives. The adoption of telework has exhibited a gradual rise, with the percentage of workers who adopt telework as the primary arrangement increasing from 4% to 6% from 2009 until the pre-pandemic period in 2020 [49]. Moreover, a larger proportion of workers engage in partial telework. The 2019 American Time Use Survey indicates that around 24% of workers do telework occasionally [196]. On the other hand, the growth in teleservice adoption is more significant. Taking online shopping, for instance, online shopping has seen a surge in popularity, with approximately 76% of American adults participating in online shopping in 2018 [216]. As the COVID-19 pandemic broke out in 2020, such a trend was completely disrupted. Following the issuance of stay-at-home

orders by many U.S. state and local governments in March 2020, a significant number of people commenced teleworking [164]. In response to the pandemic, major tech companies also implemented work-from-home policies permanently [69], though some have recently begun requesting employees to return to the office [38]. Data from a survey conducted by Pew Research indicated a substantial increase in the telework rate, rising from 20% before the pandemic to 71% during the pandemic [212]. As for teleservice, [239] showed that there was a significant surge in online shopping during the pandemic in Chicago. Their findings revealed that 45% and 67% of Chicago residents became new tele-shoppers for fresh food and grocery shopping respectively. [20] also reported that the pandemic caused around 76%-80% disruptions in out-of-home activities, such as social interactions, dine-in experiences, and shopping.

The changing trend in Tele-X adoption before, during, and after the pandemic has been shown to have fundamental impacts on individuals' activity patterns and travel behaviors, consequently influencing vehicle-miles-traveled (VMT) and traffic-related greenhouse gases (GHG) and air pollutant emissions. [238] revealed that telework may reduce commuting trips, while teleworkers are more likely to conduct more teleservices (i.e., online shopping). As a result of large-scale telework during the pandemic, both traffic congestion and emissions experienced a record drop [187]. Air quality was also found to have significantly improved due to reduced vehicle activity in March-April 2020 [150]. Nonetheless, with the economy starting to reopen in summer 2020, emissions returned to levels similar to those of 2019 in the second half of 2020 [250]. Furthermore, since people's travel behaviors changed substantially during the pandemic (e.g., avoiding transit), congestion after the pandemic might be more severe than in the pre-pandemic era [259], potentially leading to higher VMT and GHG and air pollutant emissions. Therefore, it is crucial to comprehend the trends in Tele-X adoption in the post-pandemic era, understand the resulting changes in activity patterns and travel behaviors, and evaluate their overall impacts on VMT and traffic-related emissions.

Our study reveals insights into individual choices of Tele-X in the post-pandemic era and

proposes a modeling framework to fulfill the research needs for evaluating the impacts of Tele-X. The findings from this study aim to provide policymakers with a deeper understanding of the changing trends of Tele-X and the associated impacts, enabling them to make informed decisions on how to effectively leverage its potential benefits for society in the post-pandemic era. The major contributions of this study are listed below:

- **Development of a comprehensive modeling framework:** This research proposes a comprehensive modeling framework that integrates transportation demand and supply models with air quality models to evaluate the impacts of Tele-X on both transportation systems and the environment.
- **Incorporation of activity-based demand model and agent-based simulation model:** The activity-based model (ABM) fundamentally captures individuals' travel behavior changes resulting from the adoption of Tele-X in the post-pandemic era, while the agent-based model simulates the dynamic movement of individuals and vehicles within a large-scale multi-modal network. This approach allows for the estimation of traffic volumes and traffic-related emissions with high spatial-temporal resolution.
- **Evaluation of Tele-X Impacts in Los Angeles County:** The research assesses the effects of Tele-X adoption on VMT and traffic-related emissions in Los Angeles (LA) County. Furthermore, a specific case study is conducted for a major freeway (I-405) corridor, providing a detailed analysis of the impacts on a local community.

The rest of the paper is organized as follows. Section 3.2 reviews the literature and summarizes the research gaps. Section 3.3 describes the methodology proposed in this study and introduces the design of scenarios. Section 3.4 presents and discusses the findings from the scenarios, and Section 3.5 concludes the paper and provides policy recommendations.

## 3.2 Related Work

The impacts of Tele-X have been extensively studied in the literature for decades. Within the existing body of research, four primary types of influences on in-person activities are identified: substitution (where Tele-X replaces in-person activities), complementary (where Tele-X generates additional travel needs), modification (where Tele-X adjusts existing travel needs), and neutrality (where Tele-X has little influence) [23, 145, 173, 204]. The first two types of influences, substitution and complementary, take a dominant share [140]. Specifically, telework and teleservice (e.g., online shopping) are usually studied independently [204].

Since telework was first proposed in the 1960s [193], the potential of telework to mitigate traffic congestion and improve air quality has captured the attention of policymakers and researchers for decades [87, 137, 170, 172]. While the intuitive assumption suggests that telework could reduce commute-related travel demand, some researchers have raised concerns about the extent of demand mitigation [172, 290]. Various factors may offset the benefits of telework regarding VMT reduction, including induced in-person activities resulting from flexible work arrangements [43, 131, 174], longer travel distances due to residential choices [36], increased dependency on personal vehicles [62], and others. [174] and [131] argue that telework may indirectly induce daily travel for other activities, using the time saved from commuting and flexible working schedules. Such induced travel demand may lead to unexpected congestion and air pollutant emissions in areas near teleworkers' home locations rather than their workplaces. [43] also found that people tend to conduct more in-person activities with time saved by telecommuting. Regarding travel distances, [36] found that teleworkers opt to reside farther away from downtown for more spacious and comfortable housing options. This shift in behavior may result in longer trip lengths and a higher reliance on personal vehicles, which would neutralize the benefits of VMT reduction from telework. Furthermore, a study conducted by [62] reveals that teleworking may increase weekly miles traveled, particularly by increasing car use for travel. It is important to note that these studies use data from the

pre-pandemic period. As people’s perceptions and choices of telework changed significantly after the pandemic, as well as the policies of governments and employers, it is necessary to understand people’s choices of telework adoption in the post-pandemic era and estimate the implications for travel behaviors.

As for teleservices, teleworkers’ schedule flexibility may induce additional in-person activities, while teleservices themselves can also lead to the replacement of certain physical trips, resulting in a reduction in corresponding travel demand. For instance, online shopping has been shown to reduce the number of in-person shopping trips [264], however, the overall impacts on travel demand remain unclear [40]. Researchers in public health also suggest that telehealth has the potential to reduce travel demand for healthcare purposes [280] and decrease the amount of time spent traveling to healthcare facilities [112]. Furthermore, the choices of teleservices are highly influenced by the choices of telework. [214] attempt to jointly analyze the relationship between telework, in-person shopping trips, and online shopping using pre-pandemic 2017 National Household Travel Survey data and during-pandemic travel data. The results suggest that online shopping is associated with fewer in-person shopping trips, while telework shows the opposite trend. The pandemic has even accentuated these correlations. The limitation is that only one type of teleservice (online shopping) is considered, while many types of activities exist under the category of teleservices. Further investigations encompassing various teleservices are warranted to gain a comprehensive understanding of their impact on travel behavior.

The literature review identifies two primary gaps related to the overall impacts of Tele-X. First, existing research largely focuses on the pre- and during-pandemic periods, constrained by data availability, leaving a critical need to analyze post-pandemic shifts in activity and travel behavior. Our study addresses this by evaluating how these behavioral changes affect travel patterns in a post-pandemic context. Second, there is a lack of comprehensive analyses that consider the interconnected nature of telework and teleservices alongside daily activity schedules and space-time constraints. Our research unravels these complex relationships and

provides a holistic assessment of Tele-X’s overall impacts on travel and activity behaviors.

### 3.3 Methodology

To understand the impacts of Tele-X on urban transportation systems in the post-pandemic era, we adopted a comprehensive approach that integrates a survey, an activity-based travel demand model, agent-based transportation simulation, and an emission estimation model. The framework of the methodology is presented in Fig 3.1. First, we conducted a comprehensive mixed-method survey encompassing both stated and revealed preferences in LA County to gather data on individuals’ Tele-X choices and the corresponding alterations in travel behavior. Given the notable shift in attitudes towards Tele-X following the pandemic, an up-to-date survey is necessary to grasp the nuances of these behavioral changes and enhance the calibration of the extant travel demand model. Next, we modeled the Tele-X choices and re-calibrated requisite behavioral models within the Activity-Based Model (ABM) framework to accommodate emerging travel behaviors. The disaggregated feature of ABM affords the ability to assimilate individual-level travel behavior changes, surpassing the capabilities of conventional trip-based models that are widely adopted by public agencies in the United States. Lastly, we employed an agent-based mesoscopic transportation simulation model (LA-Sim) alongside the Emission Factor (EMFAC) model from the California Air Resources Board (CARB) to assess the effects of evolving travel demand on the transportation system and traffic-related emissions. The preference for an agent-based simulation model over a static traffic assignment (STA) model stemmed from the latter’s inadequacy in addressing saturated traffic conditions [269]. Various scenarios were designed and implemented within the integrated model to evaluate the efficacy of promoting Tele-X in LA County and derive pertinent policy recommendations. The rest of this section describes the details of the methodology and model specification.

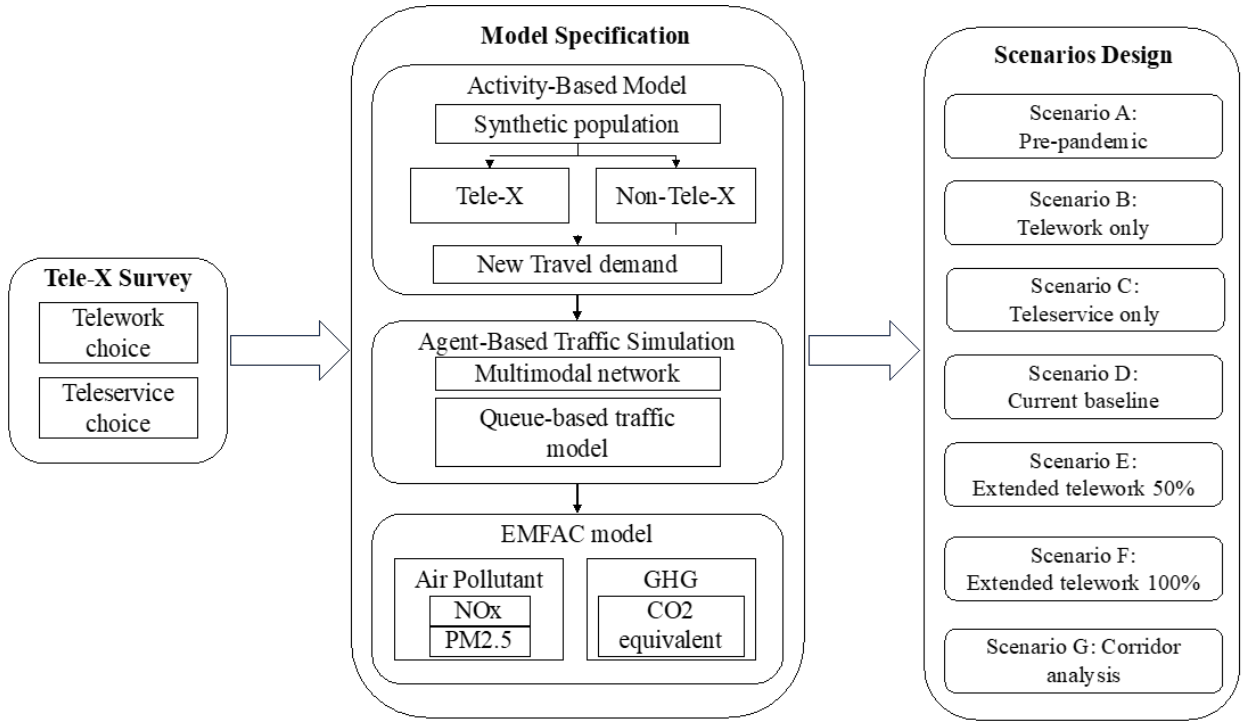


Figure 3.1: Methodology framework

### 3.3.1 Tele-X Survey

#### General Design

A mixed-method survey was designed and conducted from May 2022 to June 2022 in LA County. The survey was distributed online in collaboration with a professional survey online sampling and data collection company—Dynata [71]. Questions were asked about people’s choice of Tele-X, as well as the associated travel behavior changes. Around 1000 responses were collected, which is a reasonable size to represent the travel behavior changes [100, 117]. Significant socio-economic and demographic attributes were selected as quotas: age, ethnicity, education attainment, employment status, household size, and annual household income.

The survey consists of three layers. The first layer asks questions about the socio-economic characteristics of people. In the second layer, questions are about pre-/post-pandemic telework statuses, commuting cost and time with typical transportation modes. For those who work



remotely, we also asked about their weekly telework frequencies. In the third layer, we asked about people’s weekly activity frequencies in terms of both teleservice activities and in-person activities during the post-pandemic era.

## Survey Data Post-Processing

In total, we collected 1089 valid survey responses across LA County. Due to the limitation of online surveys, the raw survey data have some derivations in some demographic attributes. For example, the proportion of mid-age (40-65) population is lower than the quotas from the LA population, and the proportion of people with high education attainment (master and above) is higher.

To reduce bias resulting from such derivations, a raking adjustment was adopted to adjust the sample weight of each respondent and align the weighted sample with the aggregated socio-economic distributions of the population from the 2016 ABM from the Southern California Association of Governments (SCAG). The raking procedure was based on an iterative proportional fitting procedure and involved simultaneous ratio adjustments to two or more marginal distributions of population counts [126].

The raking procedure was performed in a sequence of adjustments. Base weights (sampling weights) were first adjusted to one marginal distribution and then to the second marginal distribution, and so on. One sequence of adjustments to the marginal distributions is known as a cycle or iteration. The procedure was repeated until convergence was achieved. The comparison of the weighted sample and quotas from the SCAG population is presented in Table 3.1. The demographic analysis of survey data is presented in Supplementary Information.

Table 3.1: Raw and weighted survey data compared with quotas

<b>Demographic Attribute</b>		<b>Quotas</b>	<b>Raw Survey Data</b>	<b>Weighted Survey Data</b>
1. Age	age <18	3%	1%	4%
	18<=age<25	13%	23%	13%
	25<=age<40	27%	34%	28%
	40<=age<65	40%	28%	40%
	age >=65	16%	14%	16%
2. Ethnicity	Hispanic	44%	46%	43%
	Non-Hispanic white	29%	35%	30%
	Non-Hispanic black	8%	6%	8%
	Non-Hispanic American Indian	0%	1%	0%
	Non-Hispanic Asian	16%	7%	16%
	Non-Hispanic Other	2%	4%	3%
3. Education Attainment	Less than high school	25%	5%	22%
	High school diploma	20%	33%	21%
	Associate/2-year College degree	29%	18%	30%
	Bachelor degree	17%	25%	18%
	Master's degree +	9%	19%	9%
4. Employment Status	Employed	56%	61%	56%
	Unemployed	44%	39%	44%
5. Household Annual Income	< \$25k	26%	19%	26%
	\$25k to \$50k	22%	21%	22%
	\$50k to \$75k	16%	18%	16%
	\$75k to \$100k	11%	14%	11%
	\$100k to \$125k	8%	9%	8%
	\$125k to \$150k	5%	7%	5%
	\$150k to \$200k	6%	6%	6%
	> \$200k	7%	6%	6%

### 3.3.2 Activity-Based Travel Demand Modeling

The ABM used in this study is built upon the existing modeling platform from SCAG, which is one of the largest Metropolitan Planning Organizations in the United States. It encompasses six counties in California with a population of more than 19 million. The 2016 ABM developed by SCAG is one of the largest ABMs in practice and provides a foundation for SCAG's decision-making in transportation planning. More discussions about the SCAG ABM details can be found in [100], as well as in the Supplementary Information. The SCAG ABM has the following characteristics:

- 24-hour travel demand patterns with necessary levels of temporal resolution (15-minute time intervals)
- Detailed synthetic population with demographic and socio-economic information
- Individual travel choices considering intra-household interactions

To incorporate people's travel behavior changes due to the emergence of Tele-X, several components of the SCAG ABM were updated. In the long-term choice layer, we added a telework user choice module to identify teleworkers from the workers. Additionally, we inserted a teleservice user choice module in the long-term choice layer to filter teleservice users from the general population. The selected teleworkers' workplace types were updated to "home" correspondingly in the work arrangement module. For those identified as teleservice users, updates were made in their daily activity generation layer. Specifically, their non-mandatory activity frequencies were updated according to the activity frequency choice model developed using the survey data. Similar updates were made to the maintenance and individual discretionary activity generation models. The details of the model updates are introduced in the following sub-section.

## Telework Model

This model is used to determine whether a worker in the synthetic population chooses to telework or not on a typical workday. In the survey, we designed two revealed preference questions related to workers' current telework status:

1. "Are you currently teleworking at least occasionally?"
2. "How many days per week on average are you currently teleworking?"

According to the responses to the first revealed preference question, 43% of workers telework at least occasionally. Based on the results from question 2, however, we discovered that only 17% of teleworkers telework five days a week, which aligns with the findings from the 2019 American Time Use Survey [196]. Therefore, instead of directly using the response from the first question, it is better to first develop the weekly telework frequency model using data from the second revealed preference question and then convert the weekly frequency to the equivalent daily telework frequency.

The utility functions are defined in Equation 3.1. The weekly telework frequency model has six alternatives, ranging from zero to five. Each alternative indicates the number of days a worker chooses to telework in a week. The choice of zero weekly telework frequency is regarded as the reference.

$$U_{i,n} = h(X_i; \beta_n) + \epsilon_i \quad (3.1)$$

where  $U_{i,n}$  represents the utility of individual  $i$  who chooses to telework  $n$  days a week ( $n \in [0, 5]$ ),  $X_i$  is the vector of explanatory variables for individual  $i$ ,  $\beta_n$  is the parameter vector of alternative  $n$  to be estimated from data,  $\epsilon_i$  is the random disturbance term, and  $h(X_i; \beta_n)$  is the function to calculate the systematic utility of choosing to telework  $n$  days a week. The most common specification of  $h(X_i; \beta_n)$  is a linear format:

$$h(X_i; \beta_n) = \beta_{n,0} + \sum_{k=1}^K \beta_k \cdot x_{nk} \quad (3.2)$$

where  $\beta_{n,0}$  is the alternative specific constant term,  $\beta_k$  is the parameter to be estimated for the  $k$ -th explanatory variable, and  $x_{nk}$  is the value of the  $k$ -th explanatory variable for individual  $i$ .

Assuming the random disturbance terms  $\epsilon_i$  are independent and identically distributed (i.i.d.) and follow an extreme value distribution (Gumbel distribution), the probability of individual  $i$  choosing a weekly telework frequency of  $n$  is:

$$P_i(n) = P_i \left( U_{i,n} \geq \max_{m=0,\dots,5,m \neq n} U_{i,m} \right) \quad (3.3)$$

where  $m$  denotes an alternative other than  $n$ . According to the properties of the Gumbel distribution, Equation 3.3 can be further expressed as:

$$P_i(n) = \frac{e^{\mu h(X_i; \beta_n)}}{\sum_{m=0}^5 e^{\mu h(X_i; \beta_m)}} \quad (3.4)$$

The log-likelihood across the sample is calculated as in Eqs. (3.5)–(3.6). The parameter vector  $\beta_n$  can be estimated by any optimization algorithm by maximizing the log-likelihood. We adopted the Python Biogeme package [25] for behavior model estimation, and the model estimation results are presented in Table 3.2.

$$LL = \log \prod_{i=1}^I \prod_{n=0}^5 P_i(n)^{y_{n,i}} \quad (3.5)$$

$$y_{n,i} = \begin{cases} 1, & \text{if individual } i \text{ selects alternative } n \\ 0, & \text{otherwise} \end{cases}, n \in \{0, 1, 2, 3, 4, 5\} \quad (3.6)$$

The weekly telework frequency model can be leveraged to determine the number of days that a worker chooses to telework during a week. For workers who choose to telework at least once a week, the next step is to transfer their weekly telework choices into an equivalent daily choice. The probability that a teleworker  $i$  chooses to telework on a typical workday

(any day between Monday and Friday) is given by:

$$P_i = \frac{F_i}{F_{\max}} \quad (3.7)$$

where  $P_i$  is the probability that teleworker  $i$  chooses to telework on a typical day,  $F_i$  refers to the weekly telework frequency of individual  $i$ , and  $F_{\max}$  is the maximum number of telework days, i.e., five in this case.

In our model, we first apply the telework weekly frequency model to the workers in the synthetic population, and then select a fraction of workers to be considered as teleworkers on a typical workday using the probability defined in Equation 3.7.

Table 3.2: Estimation results of telework weekly frequency model

Variable Number	Variable Name	Weekly Telework Frequency Model Coefficient (std)
<b>Constant</b>		
1	One	-2.79 (0.44) ***
2	Two	-2.07 (0.42) ***
3	Three	-2.13 (0.42) ***
4	Four	-3.14 (0.45) ***
5	Five	-2.63(-0.43) ***
<b>Household Income</b>		
6	25k to 50k	-0.386 (0.389)
7	50k to 75k	0.834 (0.371) *
8	75k to 100k	1.1 (0.398) **
9	100k to 125k	0.603 (0.428)
10	125k to 150k	1.84 (0.486) ***
11	150k to 200k	0.558 (0.479)
12	200k +	0.521 (0.495)
<b>Industry Sector</b>		
13	Manufacturing, Wholesale	-1.22 (0.347) ***
14	Retail	-0.542 (0.332)
15	Education & Health/Social Service	-0.396 (0.304)
16	Finance, Investment, Real Estate	0.237 (0.435)
17	Arts, Entertainment, and Hospitality, Food Service	-0.581 (0.368)
18	Public Administration	-1.13 (0.555) *
<b>Education Attainment</b>		
19	Associate/2-year College	0.4 (0.243)
20	Bachelor	0.993 (0.296) ***
21	Master +	1.79 (0.427) ***
<b>Household Size</b>		
22	Size = 3	0.527 (0.24) *
-, *, **, and *** indicate statistical significance at the 0.1, 0.05, 0.01, 0.001 levels, respectively		
<b>Summary statistics of weekly telework frequency model</b>		
Number of estimated parameters: 22		
Sample size: 600		
Initial log likelihood: -1075.056		
Final log likelihood: -757.0437		
Rho-square: 0.296		

## Teleservice Model

The teleservice model is intended to explore how many people in the population would choose teleservices in their daily lives and at what frequency they would use different types of teleservices. Considering that attitudes towards teleservice may differ between teleworkers and non-teleworkers, we evaluated the differences in teleservice choices across population groups and estimated the frequency at which people use teleservices for different activity types.

### *(1) Teleservice User Choice Model*

The teleservice user choice model is a binary choice model based on the revealed preference question in the survey: “Do you currently use any type of teleservices (e.g., online shopping, online education, telehealth) at least occasionally?” This model is used to select active teleservice users from the whole population. Considering that people’s employment status might impact their choices regarding the adoption of teleservices, we divide the population into the worker group and non-worker group. For workers, we further divide them into the teleworker group and the non-teleworker group. The teleservice user choice model is then segmented into three sub-models, each estimating the proportion of teleservice users in their respective population groups.

The binary logit model format is similar to the multinomial model introduced in the telework model section. The only difference is that the number of model alternatives is reduced to two: Yes or No. The estimation results of the segmented models are presented in Table 3.3.

To validate the effectiveness of the model segmentation (teleworker, non-teleworker, non-worker) regarding the teleservice choice, we conduct a likelihood ratio test between the segmented models and the pooled model. The pooled model assumes there is no difference between the three population groups regarding the teleservice choice and is estimated with

Table 3.3: Estimation results of Teleservice user choice model for teleworkers, non-teleworkers, non-workers, and pooled

Variable Number	Variable Name	Teleworkers Coefficient (std)	Non-Teleworkers Coefficient (std)	Non-Workers Coefficient (std)	Pooled Coefficient (std)
<b>Constant</b>					
1	Teleservice user	-2.62(0.52) ***	-1.88 (0.42) ***	-2.17 (0.28) ***	-2.66 (0.55) ***
<b>Housing Tenure</b>					
2	Owned	-1.43 (0.57) *	-	-	-
3	Rented	-2.1 (0.57) ***	-	-	-
<b>Age</b>					
4	18~25	-	-	-	0.99 (0.54) .
5	25~39	-	-1.43 (0.57) *	-	1.26 (0.53) *
6	39~65	-	-	-	1.07 (0.54) *
7	>65	-	-2.1 (0.57) ***	-0.754 (0.28) ***	0.47 (0.57)
<b>Gender</b>					
8	Female	-	-1.22 (0.64) .	-1.22 (0.64) .	0.53 (0.15) ***
<b>Education Attainment</b>					
9	High school	-	-	1.15 (0.31) ***	0.47 (0.24) *
10	Associate/2-year College	0.971 (0.50) .	-	1.57 (0.33) ***	0.98 (0.23) ***
11	Bachelor	-	-	1.81 (0.41) ***	1.46 (0.27) ***
12	Master +	-	-	1.56 (0.6) ***	1.54 (0.35) ***
<b>Working Status</b>					
13	Worker	-	-	-	-0.26 (0.19)
<b>Household Income</b>					
14	25k to 50k	-	0.458 (0.30)	0.494 (0.27) .	0.55 (0.18) **
15	50k to 75k	-	0.767 (0.33) *	0.652 (0.35) .	0.58 (0.12) **
16	125k to 150k	-	1.63 (0.68) *	-	0.57 (0.36)
17	150k to 200k	-	3.71 (2.07) .	0.757 (88)	0.91 (0.36) *
<b>Industry Sector</b>					
18	Retail	1.14 (0.61) .	-	-	1.18 (0.36) ***
19	Manufacturing,	-	1.11 (0.42) ***	-	0.43 (0.25) .
20	Wholesale Education & Health/Social Service	-0.602 (0.48)	0.776 (0.36) *	-	-
21	FIRE (Finance, Investment, Real Estate)	-	-	-	0.48 (0.47)
22	Public Administration	-	-	-	0.72 (0.43) .
. *, **, and *** indicate statistical significance at the 0.1, 0.05, 0.01, 0.001 levels, respectively					
<b>Summary statistics of Teleservice user choice model</b>	Number of estimated parameters: 7	Number of estimated parameters: 13	Number of estimated parameters: 14	Number of estimated parameters: 22	
	Sample size: 316	Sample size: 342	Sample size: 430	Sample size: 1088	
	Initial log likelihood: -168.06	Initial log likelihood: -251.46	Initial log likelihood: -334.97	Initial log likelihood: -754.84	
	Final log likelihood: -95.43	Final log likelihood: -204.22	Final log likelihood: -259.96	Final log likelihood: -615.65	
	Rho-square: 0.432	Rho-square: 0.188	Rho-square: 0.224	Rho-square: 0.184	



the whole data set. The likelihood ratio test statistic is calculated by Equation 3.8:

$$LR = 2(LL(\beta_t) + LL(\beta_{nt}) + LL(\beta_{nw}) - LL(\beta_p)) \quad (3.8)$$

where  $LL(\beta_t)$ ,  $LL(\beta_{nt})$ ,  $LL(\beta_{nw})$ , and  $LL(\beta_p)$  are the log-likelihoods for the teleworker model, the non-teleworker model, the non-worker model, and the pooled model, respectively. The degree of freedom  $d$  is calculated as follows:

$$d = K_t + K_{nt} + K_{nw} - K_p \quad (3.9)$$

where  $K_t$ ,  $K_{nt}$ ,  $K_{nw}$ , and  $K_p$  are the number of coefficients in the teleworker model, the non-teleworker model, the non-worker model, and the pooled model, respectively.

From Table 3.3, we can calculate the likelihood ratio test statistic as:

$$LR = 2(LL(\beta_t) + LL(\beta_{nt}) + LL(\beta_{nw}) - LL(\beta_p)) = 111.99 \quad (3.10)$$

which exceeds the critical value of  $\chi_d^2 = 32.91$  at the significance level  $p < 0.001$ . This result indicates that the three segmented models provide a significantly better fit to the data than the pooled model. Therefore, our initial conclusion is that the teleservice choice should be segmented into three models to estimate the teleservice choice for teleworkers, non-teleworkers, and non-workers.

### *(2) Teleservice Activity Frequency Model*

After we estimate people's adoption of teleservices, the next step is to estimate the frequency of teleservices used on a typical weekday. From a survey data collection perspective, it's difficult to directly ask respondents how many teleservices they use on a typical day since people's daily activity frequency may vary from day to day within a week. However, people's weekly activity pattern is usually more consistent from week to week. Therefore, by asking about respondents' weekly activity frequency, we are more likely to obtain an accurate

estimate of their number of activities under different activity purposes.

From a travel demand modeling perspective, we are more interested in investigating people’s activity patterns on a typical weekday. In this section, we follow a similar modeling approach to that described in the telework modeling section. In modeling teleservice activity frequencies, we first estimate people’s weekly teleservice frequencies by teleservice type and then convert the weekly teleservice frequency into daily teleservice frequencies.

The activities that can be conducted in the form of teleservices in the SCAG ABM include eating out (food delivery), personal maintenance (online shopping, tele-health care, online banking, etc.), and personal discretionary (online movies, online religious activities, etc.). The model estimation results can be found in the Supplementary Information. Moreover, because the alternatives defined in SCAG’s choice models for the aforementioned activity types reflect activity frequencies on a typical weekday, we need to map the weekly choice results generated from the survey data to daily frequency results as used in SCAG ABM. The detailed data process can also be found in the Supplementary Information.

### **Agent-Based Transportation Simulation**

We adopted the agent-based simulation toolkit, Multi-Agent Transport Simulation (MATSim), to simulate the movement of travelers and vehicles in a multimodal network of LA County. The modeling framework of the LA-Sim model is presented in Fig 3.2. The LA-Sim model integrates inputs comprising a multimodal network and an initial individual-level travel demand. The Mobility Simulation module simulates the explicit travel behaviors of agents within the multimodal network. A comprehensive scoring system is employed to quantify the performance of agents’ plans based on the simulated traffic dynamics. The utilities of conducting activities (usually positive) and traveling (usually negative) are both calculated. Agents possess the flexibility to iteratively refine their travel plans to enhance their daily score, according to their experience in the previous iteration. MATSim employs a co-evolutionary algorithm to ascertain the user equilibrium of the system, with scores converging after a

sufficient number of iterations. Beyond the converged simulation outcomes, we undertake an iterative calibration of road capacity to align simulated volumes with observed data. Brief introductions of model specification are presented in the following sections, and more information on the simulation model is presented in the Supplementary Information.

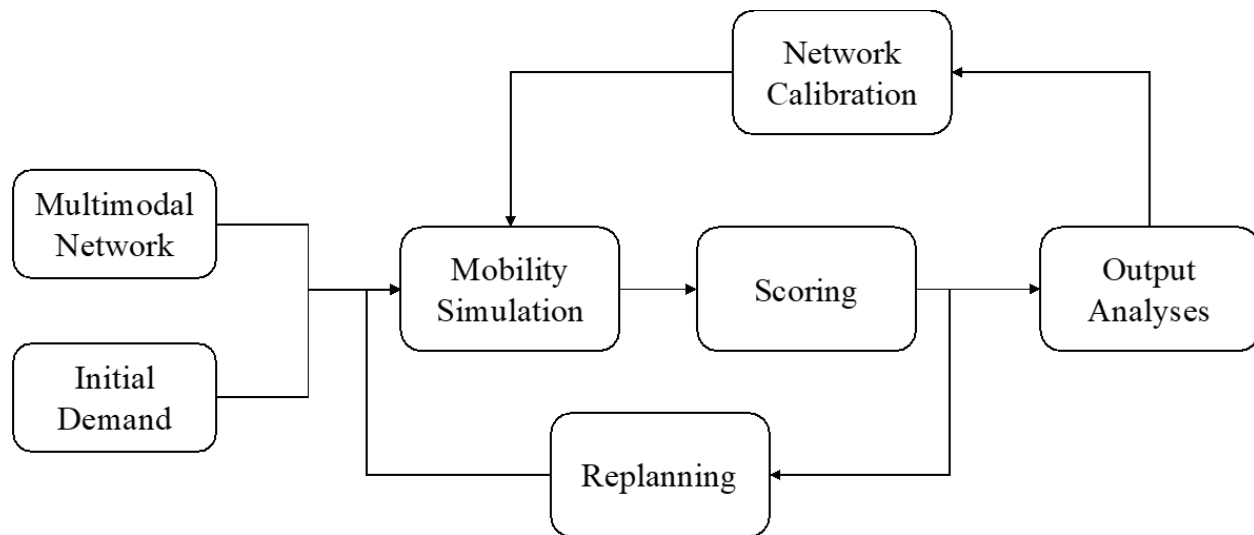


Figure 3.2: Modeling framework of the LA-Sim model

*(1) Multimodal Network*

The multimodal network consists of a road network and a public transit network, both generated from open-source datasets. The road network data was adopted from Open Street Map (OSM, 2020) for LA County, and the data was processed by the open-source Java-based network editing tool, JOSM [1], to convert it into a MATSim formatted network. The transit network was developed from the General Transit Feed Specification (GTFS, 2020) data, and we adopted the "PublicTransitMapper" class within the pt2matsim extension of MATSim to align the public transit network with the road network. Considering the scale of the LA network (354,735 links) and the low share of public transit (5%, [63]), we took an alternative route by treating transit and private vehicle usage as distinct network links to improve computational efficiency.

*(2) Initial Demand*

The initial demand was derived from the SCAG ABM with the Tele-X choice sub-models. Considering that the SCAG ABM generates the travel demand for six counties in Southern California, we filtered out daily trips for LA residents for this study. Additionally, we aggregated the origin/destination points outside of LA County to the nearest zone on the boundary so that we do not need to simulate the movement of agents outside the area of interest.

(3) *Network Calibration*

The population of LA County in 2022 is around 9.7 million [35], which is computationally expensive to simulate for the entire population in MATSim. For the LA-Sim model, we simulated 10% of the population of LA County to ensure computational efficiency. This sample simulation approach is typical in the practice of large-scale simulation, as seen in Berlin [185] and Zurich [225], which simulated 10% samples of their respective populations. Based on home locations, we randomly selected 10% of the households from the synthetic population of LA County and simulated their initial demand in the LA-Sim model. To account for the sample population, we also needed to reduce the road network capacity to some extent. However, this relationship is not linear [101]. Further calibration is necessary to capture the non-linear relationship and the impacts on road capacities from other factors, such as traffic signals and tourist trips.

---

**Algorithm 1** link capacity factor calibration

---

- 1: **Input:** link capacity factor set  $\theta^c$ ,
  - 2: **Output:** calibrated link capacity factor set  $\theta^c$
  - 3: **Initialization:** set every element in  $\theta^c$  to be 0.8 as  $\theta_k^c$ ;
  - 4: **for**  $k \in K$  **do**
  - 5:     launch simulation with  $\theta_k^c$  and calculate the  $y_{ij}^{sim}$
  - 6:     compute  $f(\theta)$  according to Equation 3.11
  - 7:     **if**  $f(\theta) < \text{threshold}$  **then**
  - 8:         stop calibration and output  $\theta_k$
  - 9:     **else**
  - 10:         update  $\theta_k^c$  according to Equation 3.12 - 3.13
  - 11:     **end if**
  - 12: **end for**
-

We proposed an iteratively proportional adjustment algorithm to calibrate the road network. The objective is defined in Equation 3.11:

$$\min_{\theta} f(\theta) = \sum_{i=1}^I \sum_{j=1}^J (y_{ij}^{sim} - y_{ij}^{obs})^2 \quad (3.11)$$

where  $\theta$  is a parameter set to be calibrated for freeway links, which includes link speed and capacity factors.  $y_{ij}^{sim}$  and  $y_{ij}^{obs}$  represent the simulated and observed volumes for traffic count station  $j$  in time period  $i$ . The link speed per hour is calibrated to the average hourly speed recorded from PeMS data [215]. The capacity adjustment factor is calibrated using Algorithm 1.

$$\theta_{(i,k)}^c = \frac{\sum_{j=1}^J y_{(ij,k-1)}^{sim} \times 10}{\sum_{j=1}^J y_{(ij,k-1)}^{obs}} \quad (3.12)$$

$$c_{(i,k)} = c_{(i,k-1)} \times f_{(i,k)} \quad (3.13)$$

where  $\theta_{(i,k)}^c$  is the capacity factor in time period  $i$  in the calibration iteration  $k$ ,  $y_{(ij,k-1)}^{sim}$  and  $y_{(ij,k-1)}^{obs}$  stand for the simulated and observed traffic volumes in time period  $i$  for count station  $j$  in calibration iteration  $k$ .  $c_{(i,k)}$  is the freeway link capacity in time period  $i$  in calibration iteration  $k$ .

## Emission Factor (EMFAC) Model

On-road emission rates of air pollutants and GHGs for Los Angeles County are retrieved from EMFAC2021 v1.0.2 [28], an official emission inventory database developed by the California Air Resource Board. We calculated a vehicle-population-weighted emission rate for each pollutant and for each emission process. Emission rates are then matched with link-level hourly vehicle volumes and vehicle activities (starting or ending a vehicle) to calculate emissions from different emission processes, including running exhaust emissions (RUNEX),

start exhaust tailpipe emissions (STREX), tire wear particulate matter emissions (PMTW), and brake wear particulate matter emissions (PMBW). The emissions from all emission processes are then aggregated to reflect the total emission of a specific link.

## Scenario Design

Using the updated travel demand model developed in this study, we created a series of scenarios to demonstrate how telework and teleservice adoption influence people’s travel choices and further impact the transportation system. These scenarios are designed to simulate demand-side variations in travel patterns at different levels of adoption for telework and teleservice, illustrating the resulting effects on transportation system performance. The scenarios are carried out at two levels: the network level and the corridor level. The network-level experiments focus on investigating the impact of telework and teleservice on the transportation system for LA County, while the corridor-level experiments concentrate on a major corridor in LA County, Interstate 405 (I-405), and analyze how changes in travel patterns in communities along I-405 influence the traffic on this corridor.

Table 3.4: Network level scenario specification summary

Index	Scenario	Model Update Specification
A	Pre-pandemic	SCAG model calibrated for 2016
B	Telework only	Implement the telework user choice model in the Pre-pandemic model
C	Teleservice only	Implement the teleservice user choice model and updated non-mandatory activity frequency model in the Pre-pandemic model
D	Current baseline	The combination of scenario B and C, accounting for current travel behaviors regarding Tele-X choices
E	Extended telework 50%	Current baseline + Increasing the telework rate in 5 industries by 50%
F	Extended telework 100%	Current baseline + Increasing the telework rate in 5 industries by 100%
G	Corridor analysis	Current baseline + Converting all non-work activities into teleservice activities in Westwood community

The base scenario incorporates both the telework model updates and the teleservice model updates into the SCAG ABM, reflecting the current telework and teleservice status of the population in LA County. To investigate the impact of telework and teleservice on the transportation system respectively, we designed two independent scenarios for telework and teleservice. One is the telework-only scenario, where only the telework model update is added to the SCAG ABM. The other is the teleservice-only scenario, where only the teleservice model update is incorporated into the SCAG ABM. Finally, two extended scenarios are designed to illustrate the effects of enhanced telework rates in five selected industry sectors. The five selected industry sectors are: Finance, Management, Professional Services, Information, and Education. These sectors are defined by McKinsey [114] as the industries with the highest potential for remote work in the United States. The specification of each scenario is presented in Table 3.4.

In the corridor-level analysis, our focus is to illustrate how further improvement in the level of teleservice in certain communities would impact traffic on major corridors based on the telework and teleservice status in the current baseline scenario. We selected I-405 as the target freeway to study how the enhanced teleservice rate would change the traffic performance on the corridor. The reason for selecting I-405 as the experiment corridor is that I-405 annually ranks as one of the most congested freeways in California and the United States. For LA County, I-405 is more than just a congested freeway—it serves as a transportation corridor of local, regional, and national significance, linking critical gateways and trade hubs. The I-405 freeway is equally important for commuters, residents, and visitors within the region. More than a quarter of LA County’s population (nearly 2.8 million residents) live within 3 miles of the I-405 freeway—known as the I-405 Corridor—and about 28 percent of jobs in LA County (1.4 million) are located within those boundaries. Please refer to the I-405 Comprehensive Multimodal Corridor Plan [167] for more detailed information about the I-405 corridor.

In this study, we concentrated on the Westwood community, which generates over 110,000

daily trips. The UCLA campus is also located in this community, and many education trips travel to UCLA. To estimate the upper bound of the teleservice impacts, we adopted an extreme assumption that all non-work activities, including school/university activities, are performed in the form of teleservice.

## 3.4 Results and Discussion

### 3.4.1 Network Experiment Results

#### Activity Analysis

Table 3.5 provides a summary of activity counts by activity types under different scenarios. The number of work activities drops in all scenarios compared to the Pre-pandemic scenario (except for the teleservice-only scenarios that have no constraint on work arrangements). The telework extended scenarios (E and F) show a massive decrease in work activities (-15.4% and -18.8%), indicating that the telework promotion policy has an effective influence on reducing work activities.

Considering the teleservice in the non-work activities, which may not result in actual trips, we divided the results into two categories: including and excluding teleservice activities. In the first category, all non-work activities are counted regardless of whether they are accomplished through teleservice. In the second category, only activities that result in actual trips are counted. The number of non-work activities, including teleservice, increases across all scenarios due to flexible activity schedules and a resulting preference for engaging in more non-work activities, except in the telework-only scenario (Scenario B). This increase is likely because higher telework rates provide greater flexibility in daily schedules, encouraging the generation of additional non-work activities. This finding is consistent with the study by Kim et al. [132], which suggests that telework tends to induce household travel.

The reason why Scenario B is an exception is that this scenario only considers the change



in telework adoption without accounting for changes in preferences regarding the adoption of teleservices. In Scenario B, the reduction in non-work activities within work tours (e.g., grocery shopping on the way home) is greater than the newly generated non-work activities due to the flexible work schedule brought by teleworking. If we exclude teleservice, the number of non-work activities barely changes, indicating that most of the induced non-work activities might be conducted online.

Let us look at Scenario D, which is the realistic scenario incorporating behavior changes in both telework and teleservice. The non-work activities (including teleservice) increased by 3.9% compared to the Pre-pandemic case, while the change in the total number of activities is only 0.2% when excluding teleservice. The decrease in work activities is offset by the increase in non-work trips.

Table 3.5: Number of activities by activity types

Index	Scenario	Work*	Non-Work (Including teleservice)	Non-Work (Excluding teleservice)	Total (Ex- cluding teleservice)
A	Pre-pandemic	$4.30 \times 10^6$	$3.83 \times 10^7$	$3.83 \times 10^7$	$4.26 \times 10^7$
B	Telework only	$3.83 \times 10^6$ (-11.0%)	$3.75 \times 10^7$ (-2.2%)	$3.75 \times 10^7$ (-2.2%)	$4.13 \times 10^7$ (-3.1%)
C	Teleservice only	$4.25 \times 10^6$ (-1.1%)	$4.08 \times 10^7$ (6.5%)	$3.98 \times 10^7$ (4.0%)	$4.41 \times 10^7$ (3.5%)
D	Current base- line	$3.79 \times 10^6$ (-11.7%)	$3.98 \times 10^7$ (3.9%)	$3.89 \times 10^7$ (1.6%)	$4.27 \times 10^7$ (0.2%)
E	Extended tele- work 50%	$3.64 \times 10^6$ (-15.4%)	$3.96 \times 10^7$ (3.2%)	$3.87 \times 10^7$ (0.9%)	$4.23 \times 10^7$ (-0.7%)
F	Extended tele- work 100%	$3.49 \times 10^6$ (-18.8%)	$3.93 \times 10^7$ (2.5%)	$3.84 \times 10^7$ (0.3%)	$4.19 \times 10^7$ (-1.7%)

## Travel Analysis

Table 3.6 shows the changes in average trip length for each scenario. We found that the average commuting distance (work trip length) decreases significantly after people shifted to telework in Scenarios B, D, E, and F. This illustrates that workers with longer commuting

distances are more likely to switch to telework than workers with shorter commuting distances. Meanwhile, the average length of non-work trips decreased in all scenarios, leading to a net decrease in the average trip length across all types of activities. Specifically, in the telework-only scenario (Scenario B), people conduct non-work trips closer to home rather than workplaces, which results in a decreased average length of non-work trips. For all other scenarios incorporating preference changes in non-work activities (C-F), people tend to engage in non-work activities near their residences when planning induced trips, contributing to the reduced trip length.

Table 3.6: Average trip length and VMT by activity type (mile)

Index	Scenario	Work		Non-Work		Total	
		Trip Length *	VMT	Trip Length	VMT	Trip Length	VMT
A	Pre-pandemic	25.96	$6.62 \times 10^7$	14.93	$2.79 \times 10^8$	16.05	$3.45 \times 10^8$
B	Telework only	24.77 (-4.6%)	$5.79 \times 10^7$ (-12.5%)	14.65 (-1.9%)	$2.77 \times 10^8$ (-0.8%)	15.58 (-2.9%)	$3.35 \times 10^8$ (-3.0%)
C	Teleservice only	25.70 (-1.0%)	$6.57 \times 10^7$ (-0.7%)	14.40 (-3.6%)	$2.79 \times 10^8$ (-0.02%)	15.49 (-3.5%)	$3.45 \times 10^8$ (-0.2%)
D	Current baseline	24.46 (-5.8%)	$5.73 \times 10^7$ (-13.4%)	14.13 (-5.4%)	$2.74 \times 10^8$ (-1.6%)	15.05 (-6.2%)	$3.32 \times 10^8$ (-3.9%)
E	Extended telework 50%	24.45 (-5.8%)	$5.53 \times 10^7$ (-16.6%)	14.11 (-5.5%)	$2.73 \times 10^8$ (-2.3%)	15.00 (-6.5%)	$3.28 \times 10^8$ (-5.0%)
F	Extended telework 100%	24.38 (-6.1%)	$5.34 \times 10^7$ (-19.3%)	14.07 (-5.8%)	$2.73 \times 10^8$ (-2.2%)	14.93 (-6.9%)	$3.27 \times 10^8$ (-5.5%)

Table 3.6 also summarizes the Vehicle Miles Traveled (VMT) by activity type for different scenarios. All scenarios show a decrease in VMT for both work and non-work trips. The realistic scenario (Scenario D) results in a 13.4% decrease in work trip VMT compared to the Pre-pandemic scenario. When the telework rate of five selected industry sectors increases to 75% and 100%, as in Scenarios E and F, the VMT drops become even more significant (-16.6% and -19.3%). From Table 3.6, we can also observe that the VMT generated by non-work trips decreases. Although the number of trips increases for most scenarios in Table 3.5, the VMT of non-work trips decreases by 1% to 2% across test scenarios. This can be explained by two factors: 1) the increased non-work trips are mostly short trips, and 2) a large portion of long-distance non-work trips are replaced by teleservice activities.

Fig 3.3 (a) provides a map showing the percentage change in VMT in the current baseline scenario compared to the Pre-pandemic scenario. The orange and red colors indicate increases in VMT, while the blue and black colors represent a drop in VMT on the links. Certain regions, such as the freeway segments north of downtown LA, the northern part of the I-405, a few east boundary links near San Bernardino County, and the southeast boundary (highlighted in red circles), show decreased VMT. This indicates that a significant portion of commute trips through these links has been replaced by telework.

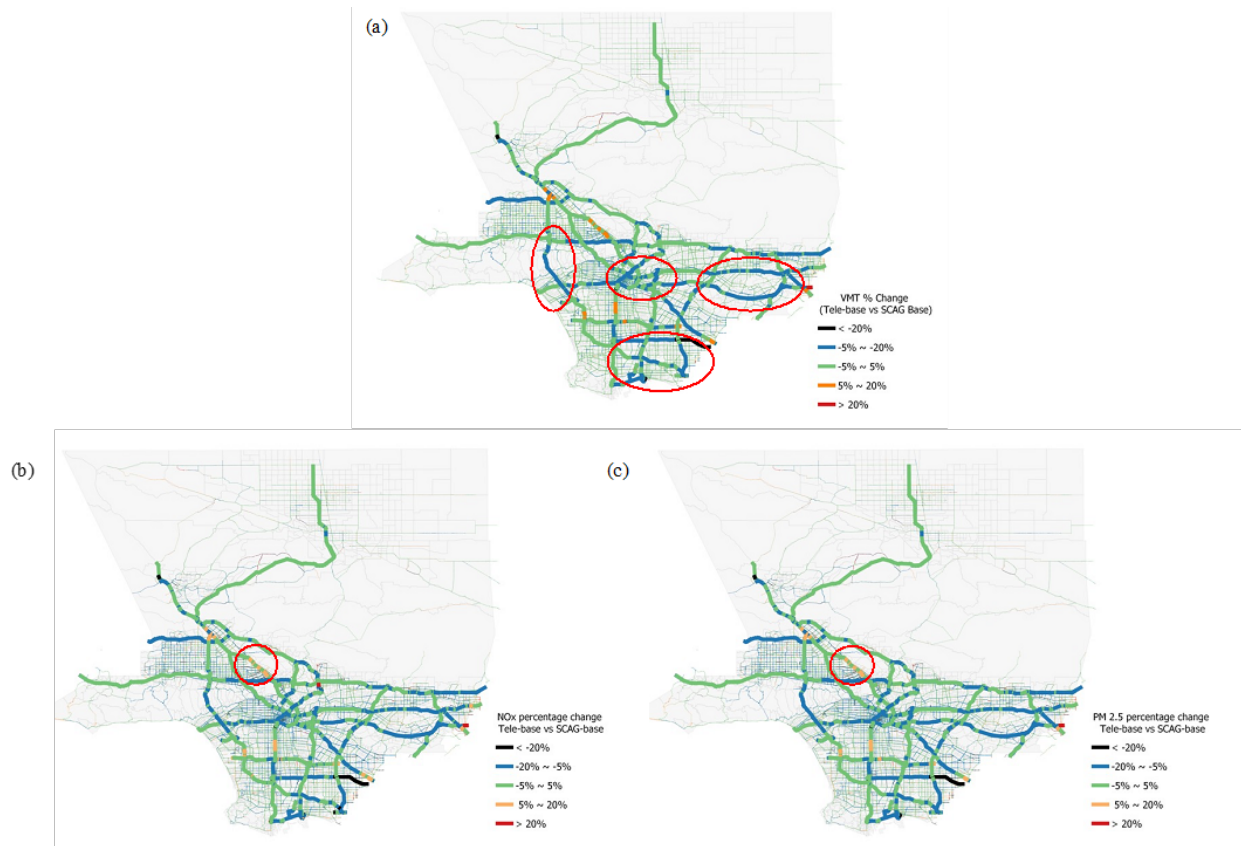


Figure 3.3: Percentage change of (a) VMT (b) NOx (c) PM 2.5 in current baseline scenario compared to the Pre-pandemic scenario (Scenario A)

## Emission Analysis

Total on-road mobile source emissions for each scenario were estimated using the latest emissions model, EMFAC22, developed by the California Air Resources Board. Although the base year of the transportation models (SCAG ABM and LA-Sim) is 2016, the emission inventory changed significantly from 2016 to 2022. Considering the travel patterns in LA County did not change drastically, with the daily VMT decreasing by around 4.4% from 2016 to 2022, we used EMFAC22 to estimate the emissions from traffic. The differences between the test scenarios (Scenarios B-F) relative to the baseline (Scenario A) indicate the relative benefits of these scenarios. The emission reduction results are provided in Table 3.7. We categorized the emission sources into two types: air pollutants (NO<sub>x</sub> and PM<sub>2.5</sub>) and greenhouse gases (CO<sub>2</sub> equivalent). The first category focuses on emissions hazardous to health with localized impacts, while the second category focuses on emissions that contribute to climate change.

Compared to the Pre-pandemic scenario (Scenario A), all test scenarios except Scenario C reveal emission reduction benefits for both air pollutants and GHG. Note that the relative reduction in VMT does not directly correspond to the reduction in emissions. Scenarios with greater VMT reductions in Table 3.6 do not necessarily result in a greater reduction in emissions. Scenarios C and D exhibit more VMT reduction than Scenario B, but they show less emission reduction than Scenario B. This can be explained by the fact that the increase in trip frequency and associated exhaust emissions from trip starts can offset the benefits of reduced VMT in terms of emission reduction.

Figs 3.3 (b) and (c) provide the percentage emission reduction along the freeways within the LA County boundary of the current baseline scenario (Scenario D) relative to the pre-pandemic scenario (Scenario A). As seen in Fig 3.3 (b) and (c), the reduction patterns in air pollutants are similar to the VMT. In some areas, however, such as the freeways north of the downtown region (highlighted in the blue circle), emissions are higher than in the baseline

Table 3.7: Emissions by pollutant type (ton per day)

Index	Scenario	Air Pollutants		GHG
		NOx*	PM2.5	CO2 equivalent
A	Pre-pandemic	62.9	2.99	$1.25 \times 10^5$
B	Telework only	61.3 (-2.6%)	2.90 (-2.9%)	$1.21 \times 10^5$ (-2.9%)
C	Teleservice only	63.8 (1.3%)	3.00 (0.6%)	$1.26 \times 10^5$ (0.6%)
D	Current baseline	61.6 (-2.1%)	2.89 (-3.1%)	$1.21 \times 10^5$ (-3.1%)
E	Extended telework 50%	61.5 (-2.3%)	22.89 (-3.2%)	$1.21 \times 10^5$ (-3.2%)
F	Extended telework 100%	61.3 (-2.6%)	2.88 (-3.6%)	$1.21 \times 10^5$ (-3.6%)

\* Count (change in % comparing to Pre-pandemic)

scenario. This occurs because, for these regions, the VMT increment caused by the growth of non-work trips exceeds the VMT reduction induced by the decrease in work trips.

### 3.4.2 Corridor Experiment Results

As introduced in the experiment specification, we selected the Westwood community near the I-405 to illustrate how changes in activity patterns impact traffic and emissions on the I-405. All the university/school and non-mandatory activities (eating out, personal/household maintenance, discretionary, etc.) are converted to teleservice as an extreme case. For the aforementioned trips, it is important to analyze where they originate and how the elimination of these trips affects the I-405 corridor. Fig. 3.4 shows the origins of the trips with specific activity types that end in the four selected communities. Note that some boundary traffic analysis zones (TAZs) show relatively large trip origin counts. This is because trips originating from outside LA County are aggregated and assigned to the boundary TAZs adjacent to those external counties.

Fig. 3.5 (a)-(b) provide the temporal distribution of the traffic volume and travel speed on I-405 across different scenarios. Two major findings can be observed in Fig. 3.5 (a) and (b). First, the traffic volume during morning peak hours is reduced in the Westwood scenario (Scenario G) compared to the Pre-pandemic scenario (Scenario A). Specifically, the volume

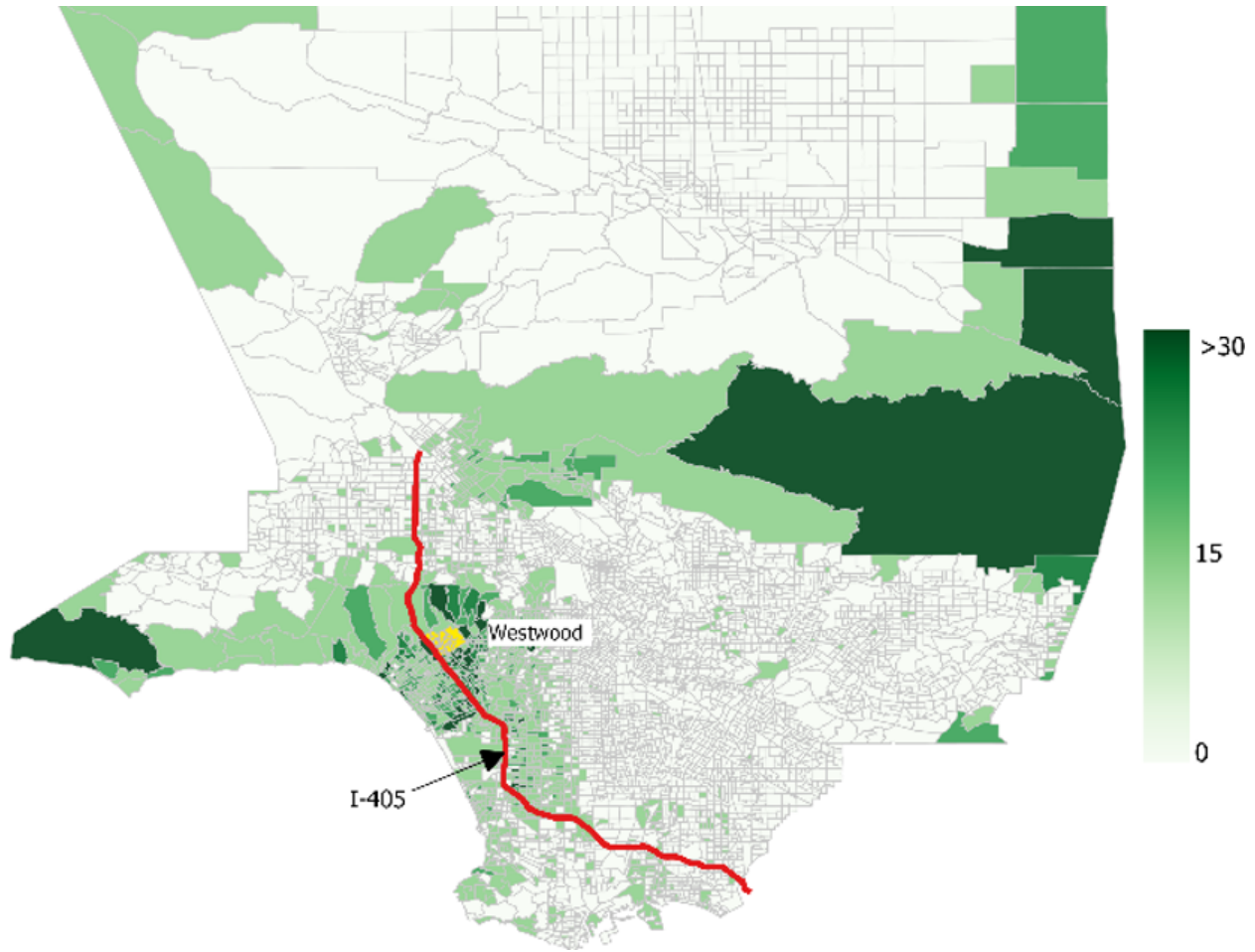


Figure 3.4: Spatial distribution of non-work trip origins from elsewhere, ending in Westwood communities

from 6 to 7 AM decreases from 4.63 million to 4.27 million (-9%). This reduction is due to the decreased work trips caused by the increased adoption of telework in Scenario G, which is similar to the current baseline scenario (Scenario D). Second, the reduction in school and non-mandatory trips in the Westwood area leads to a significant increase in traffic volume on I-405 between 4 PM and 10 PM, compared to the other two scenarios (Scenarios A and D). This can be explained by the reduced trips to Westwood, which decreased congestion on I-405, thus attracting more vehicles to use the freeway. As for the travel speed, the average distribution does not change significantly for I-405.

However, Fig. 3.4 indicates that trips to Westwood are highly concentrated along the

I-405 corridor, making these trips heavily reliant on this freeway segment. The impacts on the freeway segment in the Westwood region are more pronounced. Fig. 3.5 (c)-(d) provide a more detailed spatial analysis by showing the traffic volume on the freeway segment within the Westwood region.

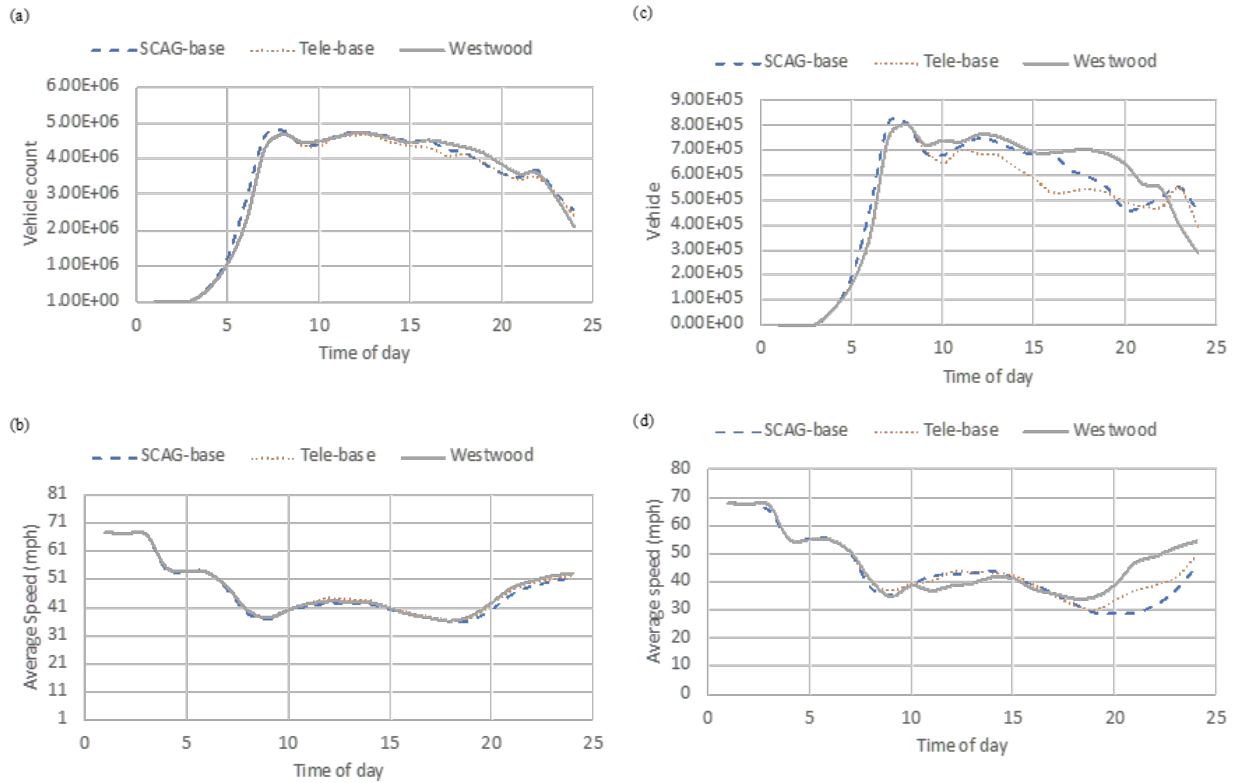


Figure 3.5: Temporal distributions of (a) traffic volume and (b) travel speed on I-405 and (c) traffic volume and (d) travel speed on Westwood segment

As seen in Fig. 3.5 (c) and (d), the traffic volume increase due to the trip reduction in Westwood is more pronounced. The current baseline scenario (Scenario D) significantly reduces traffic volume on the freeway segment during the day (6 AM – 7 PM) compared to the Pre-pandemic scenario (Scenario A), while the traffic volume increases in the Westwood scenario (Scenario G) from 5 PM – 10 PM, exceeding even the Pre-pandemic scenario (Scenario A). This finding can be attributed to the extreme assumption of non-work trip cancellation, which reduced local trips that would have used the I-405 segment, inducing



more traffic on the segment and increasing traffic volume.

If we compare the distribution of traffic volume and travel speed, another interesting finding emerges. During the period (5 PM - 10 PM) when the traffic volume increases in Scenario G, the travel speed on the I-405 segment also increases. This result appears counterintuitive, as it is generally expected that travel speed would decrease when traffic volume rises on a congested freeway segment. One possible explanation for this unexpected finding is the extreme non-work trip cancellation assumption, which effectively removed local trips that frequently use on-ramps and off-ramps. By eliminating the congestion and delays caused by on-ramp/off-ramp traffic, the segment's travel speed increases despite the overall volume increase.

Besides the network-level emission analysis in the previous section, we also conducted a corridor-level emission analysis to investigate the emission impact on the Westwood community compared to the Pre-pandemic scenario. Table 3.8 provides a summary of the emissions divided by emission types. As indicated in Table 3.8, the current baseline scenario (Scenario D) demonstrates a significant reduction in emissions, with approximately -7.2% to -7.3% reduction in NO<sub>x</sub> and PM<sub>2.5</sub> on the I-405 segment, signifying a notable achievement in emission reduction. However, when we further promote teleservice specifically in the Westwood community (Scenario G), air pollutant emissions unexpectedly increase by 2.2% and 2.1%, respectively.

This finding suggests that Tele-X could indeed serve as an effective solution to alleviate traffic congestion and traffic-related emissions in local communities. Nevertheless, it highlights the need for cautious consideration when determining the extent of Tele-X promotion. Transportation systems are intricately connected and dynamic, meaning that promoting Tele-X in one community might temporarily alleviate the situation but could lead to an influx of traffic from other areas, exacerbating the overall situation. Consequently, a comprehensive and balanced approach is needed to implement Tele-X strategies in various communities, accounting for potential spillover effects and ensuring that the maximum benefits for traffic



congestion and emissions reduction are achieved.

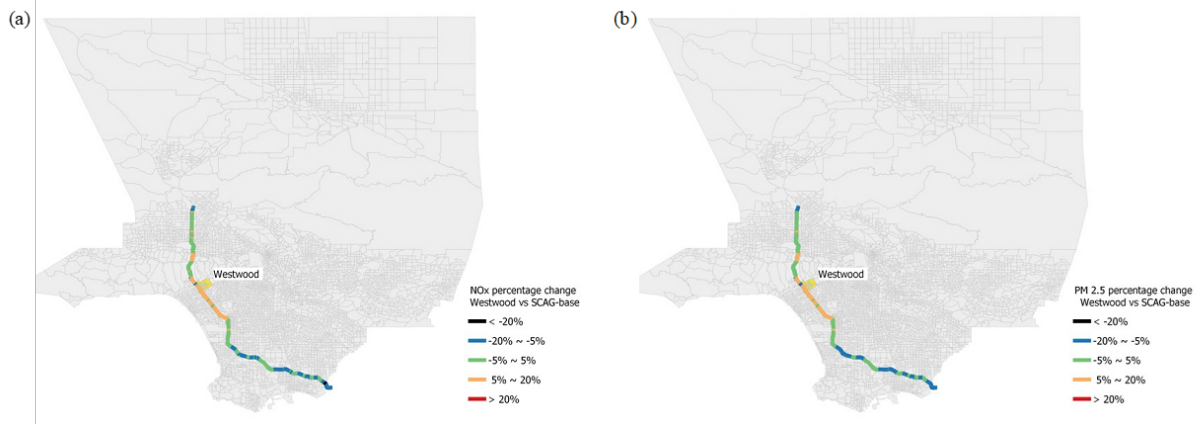


Figure 3.6: NOx and PM 2.5 percentage change relative to Pre-pandemic scenario on I-405 for Westwood

Fig. 3.6 illustrates the emission change distribution of NOx and PM2.5 along the I-405. The general trend shows that promoting teleservice is more likely to increase emissions near the central and upper segments of I-405, while reducing emissions on the southern segments of I-405. This is because a full transition to teleservice in Westwood would reduce local needs for I-405 and, therefore, attract more traffic from nearby areas. The higher passing-by traffic volume results in increased emissions for the local community.

Table 3.8: Air pollutant emissions generated on the Westwood segment of I-405 (ton per day)

Pollutant	Index	Scenario	Emission*
NOx	A	Pre-pandemic	0.409
	D	Current baseline	0.380 (-7.2%)
	G	Westwood	0.418 (2.2%)
PM 2.5	A	Pre-pandemic	0.0220
	D	Current baseline	0.0204 (-7.3%)
	G	Westwood	0.0225 (2.1%)

\* Count (change in % comparing to Pre-pandemic)

## 3.5 Conclusion

The outbreak of the COVID-19 pandemic has engendered profound repercussions across multiple facets of individuals' daily lives. The ensuing alterations in activity preferences and travel behavior, stemming from pandemic-induced impacts, are poised to persist well beyond the pandemic's cessation. This study focuses on the impact of Tele-X on people's travel patterns and the consequential shift in the transportation system and associated emissions. This research collected people's activity and travel choice data through an online survey, estimated the changes in travel demand by updating the SCAG ABM with the survey data, and further conducted an agent-based mesoscopic traffic simulation to analyze how the transportation system performs under the consideration of Tele-X. The study also designed several test scenarios to demonstrate the performance of the transportation system when telework or teleservice statuses change. Selected findings from this study that can inform policymakers about the features of post-pandemic travel and help them find solutions to enhance transportation performance include:

- The post-pandemic activity frequency experiences an increase compared to the pre-pandemic period. While work trips in the post-pandemic era decrease by 11.7% due to telework, non-work trips surge by 1.6%, yielding an overall marginal change in total trips (about 0.2% growth).
- The adoption of Tele-X results in a decrease in VMT (3.9% lower than the pre-pandemic era). While the non-work trip-induced VMT barely changes (-1.6%), the VMT drop is mainly contributed by work trips (-13.4%).
- The adoption of telework and teleservices exerts a minor impact on ground transportation-related emissions (NO<sub>x</sub> -2.1%, PM<sub>2.5</sub> -3.1%, and CO<sub>2e</sub> -3.1%). This general trend is upheld across the majority of Los Angeles County, albeit certain regions, particularly those north of downtown LA, exhibit augmented emissions due to induced non-work

trips.

- The present Tele-X trend yields a modest diminution in peak-hour traffic and perpetuates a closely analogous diurnal pattern in traffic volume and speed relative to the pre-pandemic era.
- The promotion of teleservice in communities near the central parts of I-405 (Westwood) would induce more passing-by traffic on the corridor. This is because trips to Westwood predominantly originate from nearby areas and are highly dependent on I-405. Eliminating those trips through teleservice would reduce pressure on I-405 and induce more passing-by traffic.

It should be acknowledged that while the changes in total trip count, VMT, and traffic-related emissions may not be of paramount magnitude, statistical analysis may be warranted to ascertain the degree of confidence surrounding these changes. The limitation of this study is that given the computational complexity of the proposed methodology, only a singular simulation run was executed within the scope of this research. However, the representativeness of similar practices for large-scale simulation models has been corroborated by prior studies [46]. The rigor of this research can be further bolstered by conducting multiple simulation runs in subsequent endeavors.

This study can also be improved in several ways. First, the network-level analysis can further incorporate equity considerations to assess disparities in travel across different communities. Second, the corridor-level analysis can be conducted in a more detailed approach by extracting the trajectories of each vehicle and focusing only on the trips that pass through a selected corridor. Third, people's long-term choice changes can be added to the ABM, as literature indicates that choices, including home location and vehicle ownership, may change due to telework [36].

## Chapter 4

# Gap assessment and demand projection of public charging infrastructure in electrified transportation systems

Transportation system electrification is expected to bring millions of electric vehicles (EVs) on the road within decades. Accurately predicting the charging demand is necessary to accommodate the surge in EV deployment. This study presents a novel modeling framework to predict the public charging demand profile derived from people's travel trajectories, with consideration of the demand and supply stochasticity of transportation systems and the charging behavior heterogeneity of EV users. The vehicle charging decision-making process is explicitly modeled, and the charging needs of each EV user are estimated based on their travel trajectories. The methodology enables charging demand prediction with high spatial-temporal resolution for transportation system electrification planning. A case study was conducted in Los Angeles County to predict the demand for public charging facilities in 2035 and perform a corresponding spatial-temporal analysis of EV public charging under various scenarios of future electrification levels and network conditions.

## 4.1 Introduction

The rapid development of advanced vehicle technologies (e.g., electric vehicles (EV), automated vehicles (AV)) in recent decades has brought various benefits to society, such as reducing petroleum consumption and decreasing greenhouse gas (GHG) emissions [118, 284]. Plug-in electric vehicles (PEV or EV), further classified as battery electric vehicles (BEV) and plug-in hybrid electric vehicles (PHEVs), play a significant role in reducing the nation’s dependence on fossil fuels and achieving zero tailpipe emissions [272]. To achieve climate mitigation goals, both federal and many state governments have required the acceleration of transportation sector electrification [110]. For instance, the state of California issued an executive order in September 2020 requiring that, by 2035, all new passenger cars and trucks sold in-state must be zero-emission [271]. Specifically, the city of Los Angeles plans to have 80% of total vehicles be EVs by 2035.

The adoption of EVs is essential to achieving regional electrification goals. Moreover, people’s willingness to adopt EVs has a strong correlation with access to charging infrastructure [95, 217]. As of 2019, the planned charging infrastructure in California had only reached around 40% of the target number required to support the projected fleet size of EVs by 2025. Furthermore, the spatial distribution of public chargers shows significant disparities across communities in California. [110] found that access to public chargers in lower-income communities is significantly lower than in upper-income communities, based on public charging facility data from California in 2021. The allocation planning for public EV charging infrastructure cannot be performed without understanding the potential spatial and temporal distribution of charging needs [153]. Therefore, accurately predicting charging demand is fundamental for regions aiming to meet future electrification goals. A valid estimation of charging demand, including peak periods and demand locations, would greatly support the deployment of the EV market and inform regional grid system upgrades.

### 4.1.1 Related Work

Many previous studies have explored different methods for charging demand prediction of various types of EVs, such as electric taxis [278], electric buses [10], and passenger EVs [202,262]. Considering that EV passenger cars dominate the EV market share, accounting for over 90% of all EV types [113], this study concentrates on predicting the charging demand for passenger EVs. There are two main approaches for predicting the public charging demand for passenger EVs. One method is extracting the charging demand from real data by learning the charging patterns directly from the charging diaries of existing EV fleets, while the other method derives charging needs from travel trajectories, either from travel surveys or models.

The first method utilizes actual data collected from test EV fleets or selected EV charging stations to generate the EV charging demand. [262] developed a stochastic method to generate the daily charging demand pattern for EV users based on driving trajectories and charging session data from a real-world EV fleet. [10] implemented trajectory data of real taxis and developed a model for estimating EV taxi drivers' charging patterns under the assumption of limited charging locations. [257] developed a regression model to estimate the key patterns in the use of public chargers, based on charging session data from more than 200 public charging locations in Amsterdam. [202] used Caltech's Adaptive Charging Network (ACN) data, extracted from two public charging points located on campus, to analyze the occupancy status of both the single charger and the entire charging station. These studies, however, are limited in EV fleet size, making it difficult to expand the results to a large-scale transportation network. Moreover, these studies, focused on charging data, lack the capability to comprehensively understand how charging would influence people's daily activities and travel behaviors. Therefore, a detailed travel diary comprising spatial-temporal varied travel-related choices is essential for studying the charging demand profile of a large EV fleet and its related characteristics.

Besides data-driven EV charging demand prediction, another commonly adopted method

is to derive the charging demand from people’s travel trajectories. This method requires a detailed profile containing the location and schedule information of each individual, which is typically generated either by travel surveys or transportation simulation models. Some studies focused on optimizing charger locations [67, 129], while in this study, we pay close attention to the needs of public charging facilities at the zonal level, helping decision-makers understand the charging needs distribution and plan accordingly for the investment and deployment of public charging facilities. A common approach in modeling EV charging patterns involves using regional or nationwide travel surveys to generate conventional vehicle usage patterns [59]. These methods are often applied in combination with charging behavior scenarios to create charging profiles. [107] predicted EV charging load based on data extracted from the 2009 National Household Travel Survey. They developed a Monte Carlo model from the survey data and estimated the 24-hour charging load generated by EVs. [177] modeled EV charging behavior by assuming a determined set of daily travel distances, charging durations, and charging power for EV users in Western Australia. They designed several fixed EV charging scenarios to determine the time windows for charging sessions and generated variable start times for charging events by introducing stochastic delays for each charging session. [33] predicted specific daily charging demands at various destinations and times based on data from the 2009 National Household Travel Survey on vehicle travel distance and origin-destination patterns. The authors conducted experiments to determine EVs’ charging needs, based on battery capacity and state of charge. [245] investigated the spatial and temporal distribution of charging load in Berlin by integrating travel patterns extracted from the 2017 German Household Travel Survey (GHTS) and sociodemographic characteristics such as population density, household income, and auto ownership levels from census data. A major limitation of these survey-based models is that they ignore the impacts of charging strategies on EV travel patterns.

Apart from the survey data-based method, transportation simulation modeling is another popular approach for analyzing EV charging patterns. Agent-based or activity-based modeling

(ABM) is one of the most well-accepted methods in EV charging demand applications. [133] developed an activity-based travel demand model to obtain each vehicle’s schedule and established a mapping rule to represent EVs of different battery capacities with gasoline vehicles of varying engine volumes. [85] developed an activity-based EV charging demand model for small-scale networks using the open-source traffic simulator SUMO. The authors assumed that EV users’ mobility is not affected by charging needs and that charging sessions occur at stations close to activity destinations, minimizing the impact on existing schedules. [278] designed an integrated model combining an activity-based model and the evolution of EV penetration to estimate EV charging patterns and energy load. The model also considers vehicle purchasing preferences across sociodemographic groups. [141] presented an agent-based EV model to investigate the impact of EVs on the grid and validated the model with real-world observed data. By adopting the agent-based approach, these models capture the stochasticity in daily travel for each individual in the transportation system. However, existing agent- or activity-based demand prediction models rarely consider variations in travel demand or supply, such as weekday/weekend travel differences or roadway capacity variations under different conditions. For instance, some literature finds that traffic conditions (e.g., traffic speed, number of stops) can largely affect EV energy consumption [81, 286], thereby influencing the timing and location of charging activities.

In the travel-based EV charging prediction approaches mentioned above, charging behaviors or strategies of EV users are simulated in several ways. Typical charging behaviors include: 1) *Uncontrolled/Uncoordinated charging*, where EVs start charging whenever they arrive at destinations and stop when fully charged or leaving for the next activity [85, 142]; 2) *Controlled/Coordinated charging*, where users make charging decisions to minimize costs, which may include total energy consumed, voltage deviation, current, etc. [235, 266, 286]; 3) *Delayed charging*, where charging is assumed to be delayed until the evening when the cost is minimized [56, 190]; and 4) *Off-peak charging*, where charging is directly controlled by central operators and performed only during off-peak hours [82, 93]. A significant limitation of these



charging behaviors is that they are based on predetermined charging scenarios that do not fully account for the heterogeneity in EV users' charging decisions.

### 4.1.2 Main Contribution

The primary gaps in existing research include: 1) the absence of large-scale, robustness predictions for EV public charging demand, 2) the lack of a high-resolution charging profile that captures detailed individual EV usage, and 3) limited modeling of variable travel demand and supply impacts on charging demand. To address these, we developed the CREATE framework (Charging Reliability, Resilience, Equity, and Accessibility in Transportation Electrification). The proposed framework delivers high-resolution, data-driven charging forecasts and proposes strategies for equitable and resilient charger deployment, ensuring integration with the power grid. Our framework provides a comprehensive approach to optimizing charger placement and addressing diverse challenges in transportation electrification. While CREATE is a comprehensive system comprising the entire process from charging demand generation, charging system reliability evaluation, to resiliency enhancement, this study focuses on the first goal: predicting public charging demand. The fundamental principle is that EV charging activities originate from EV users' travel activities and are influenced by variations in both travel and EV usage. This study considers factors that lead to variations in electrified transportation systems from the following aspects: 1) Travel demand-side variations, such as weekday/weekend demand differences; 2) Transportation supply-side variations, such as congestion and road incidents that may influence travel schedules; and 3) Charging behavior variations, such as triggering battery state of charge (SOC) for charging, range anxiety, charging duration, and location. this study focuses on public charger demand prediction since public charging infrastructure is a key factor influencing EV adoption from the public service perspective [152, 183]. The main contributions of the proposed framework are as follows:

- A comprehensive demand prediction modeling framework for EV public charging is

developed. The framework enables the explicit modeling of individuals' charging events derived from activity schedules, travel trajectories, and charging decision-making. The high-resolution transportation simulation improves charging demand prediction accuracy.

- The system adopts a scenario-based approach to capture demand and supply-side variations of transportation systems and incorporates their impacts on public charging demand. Combinatorial scenarios are simulated by adjusting travel demand and network supply to represent variations in seasons, days of the week, weather, special events, incidents, work zones, and more.
- A comprehensive charging decision-making module is developed to explicitly generate compulsory and conditional charging decisions from the simulated activity schedules and travel trajectories of EVs under each scenario, representing heterogeneous charging habits.
- This study investigates public charging demand and the spatial-temporal distribution of public chargers at both the network and zonal levels. The model outputs estimate public chargers and accumulative installation costs under various future government/agency electrification goals as future investment recommendations.
- The equity of public charging accessibility across different socio-demographic groups is evaluated using a statistical approach. Charging access performance for both the target year and current year is quantified to indicate how public charging equity evolves with electrification levels and projected public chargers.

## 4.2 Methodology

### 4.2.1 Model Overview

This section provides an overview of the model for public charging demand prediction in the CREATE system. The model described here is a holistic demand profile prediction system for public charging infrastructure in electrified transportation systems at the zone level, such as traffic analysis zones (TAZ) or census tracts (CT). The outline of the model with its three major components is illustrated in Fig. 4.1.

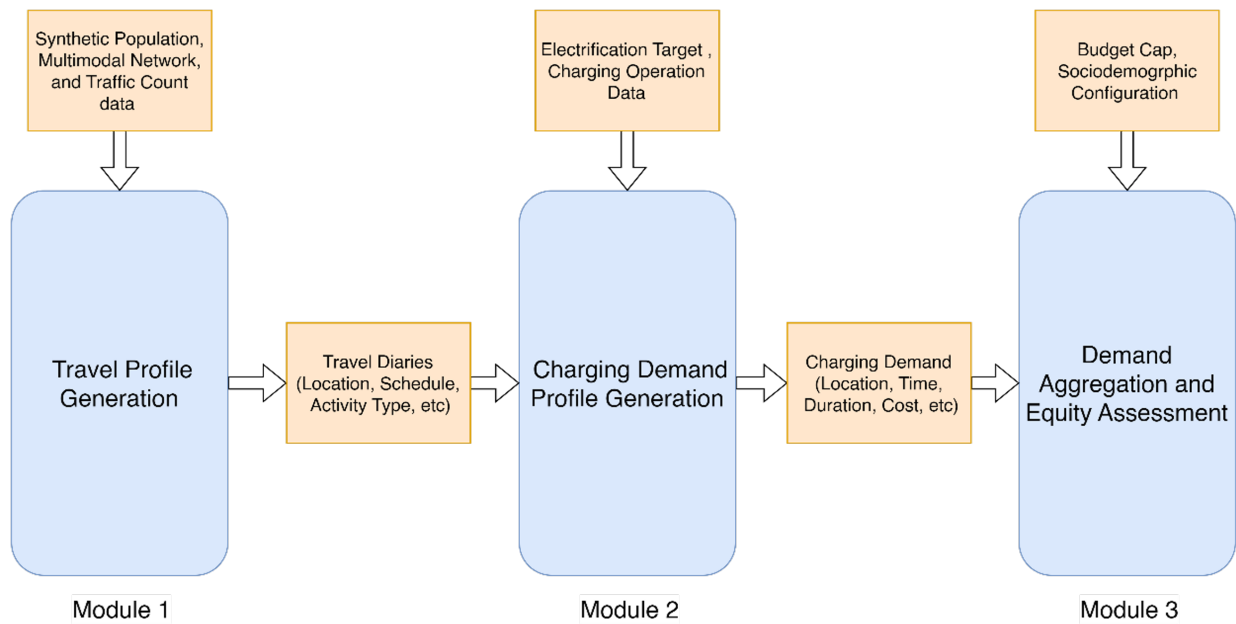


Figure 4.1: Model Framework Overview

The *travel profile generation module* is designed to create the daily travel trajectory for each individual in the target region. This module inputs synthetic population data and multimodal network data into travel demand and supply models and outputs travel trajectory data that includes origins, destinations, activities, schedules, travel modes, routes, and travel times for each agent. The key feature of this module is that it captures variations in real-life transportation operations from: 1) the travel demand side, including demand variations for

seasons, weekdays and weekends, day-to-day differences, and special event days, and 2) the traffic supply side, such as the impacts on road capacities due to weather, incidents, and work zones.

The *charging profile generation module* provides an accurate estimation of the location, timing, and duration of every single charging need occurring during a typical day. The charging demand generation is carried out under different electrified transportation operations determined by future electrification levels and initial state of charge (SOC). First, the future agency electrification level sets the target year and determines the EV fleet size and private charger distribution. Second, given each electrification level, the model classifies a typical day by different initial SOC levels: low initial SOC and high initial SOC. The initial SOC reflects the SOC level of EVs at the beginning of the day, as adopted by LA100 [50].

Using the charging demand generated from the previous module, the last model component provides an analysis of the spatial-temporal pattern of charging demand and predicts the required number of chargers given a specific investment goal. With the estimated number and location of chargers, we also conducted a charging equity analysis across different communities in the region. The charging equity performance for the target year is compared with the current year using existing charging facility data. This allows us to estimate how the disparity in charging accessibility across different socio-demographic groups may evolve.

## **4.2.2 Travel Profile Generation**

### **Activity-Based Transportation Simulation System**

The activity-based transportation simulation system is a comprehensive implementation of tools from ABM and dynamic transportation simulation. The ABM adopted in this study was developed by SCAG (Southern California Association of Governments) [102, 118], and the transportation simulation platform used is the multi-agent transportation simulation (MATSim) toolkit. Both can be replaced with other widely applied ABM and dynamic

transportation simulation platforms.

ABM is a regional transportation plan model based on individuals' activity desires, which forecasts their travel demand. Based on a synthetic population with demographic and socio-economic attributes, the ABM predicts all activities of a person as a 24-hour trip chain, connecting activities in time and space with details such as origin, destination, and scheduled start/end time of activities. The original ABM was developed to represent the travel demand of the target region on a typical weekday. To capture the variation in travel demand between weekdays and weekends, we expanded the scope of the original ABM to generate travel demand for both weekdays and weekends.

With the travel demand provided by the ABM, MATSim simulates the operation of the transportation system using additional inputs, such as a topological multimodal network and road and vehicle attributes [102]. MATSim adopts an iteratively co-evolutionary algorithm [108] to search for the user equilibrium of the system. Variations in the travel supply side, such as changes in road capacity due to incidents, extreme weather, and work zones, can also be incorporated into the simulation. The output of the simulation model is the daily travel trajectories for all individuals, presented as a tour list table containing vehicular movements and time-dependent statistics of trips and activities, accounting for the mutual influence of vehicles on roads.

## **Scenario-Specific Travel Demand and Transportation Network**

This section explains how demand variations are modeled and how various network scenarios are generated. The method of defining various travel demand and traffic network scenarios is similar to what has been adopted by SHRP 2 Reliability Project L04 [162].

First, a travel demand profile was created to account for various temporal variations, such as seasonal changes and weekday/weekend differences. The method for generating travel demand files for these variations involves adjusting the baseline travel demand output produced by the original ABM, as detailed in Section 4.2.2. Since the baseline ABM only

simulates demand for a typical weekday, it is necessary to modify the original ABM output to reflect demand across different seasons of the year and days of the week. This adjustment ensures that the temporal distribution of travel demand aligns with patterns extracted from historical traffic count data collected by detectors.

Secondly, the method for creating simulation networks with varying supply-side conditions involves adjusting network link capacities based on historical data sampled throughout the year. A mix-and-match approach is used to combine all possible external variations, from seasonal demand patterns to weather conditions. For instance, one example of a traffic network variation could involve a spring season, a Monday weekday, a non-holiday, and rainy weather. Each variation in the combination may alter traffic demand (e.g., increased demand on a holiday compared to a regular day) or network supply (e.g., reduced road capacity during rainfall compared to clear weather). The addition of one variation layer means the occurrence of this layer under all previous layers, whose probability,  $p(\text{TN}_{1,2,\dots,i})$ , satisfies the multiplication rule based on conditional probability, Equation 4.1. ‘TN’ stands for the transportation network, and the subscripts  $1, 2, \dots, i$  represent the variation layers from the season, working day, and day of the week to the weather. These layers target the entire traffic network.

$$p(\text{TN}_{1,2,\dots,i}) = p(\text{TN}_i | \text{TN}_{1,2,\dots,i-1}) \times p(\text{TN}_{1,2,\dots,i-1}) \quad (4.1)$$

The last variation layer of the scenario-specific transportation network relates to traffic incidents or work zones that focus on specific road segments in a network. This layer concerns the temporal and spatial uncertainties of network operations. Therefore, Monte Carlo simulation is well-suited to generate finite network-level traffic incidents or work zone cases. These cases are under the condition of different network variation combinations, which vary from season to weather, 1 to  $i$ . Equation 4.2 gives the probability of the last layer  $j$ , similar to Equation 4.1. The conditional occurrence probability of layer  $j$ ,  $p(\text{TN}_j | \text{TN}_{1,2,\dots,i})$ , is the

statistical distribution of the daily incident/work zone frequency based on their historical data, together with the information on incident/work zone start times and associated durations.

$$p(\text{TN}_{1,2,\dots,i,j}) = p(\text{TN}_j|\text{TN}_{1,2,\dots,i}) \times p(\text{TN}_{1,2,\dots,i}) \quad (4.2)$$

The above process implements the transportation network scenario generation with demand and supply variations, particularly in the temporal aspect. To distribute the simulated scenarios spatially into the network, the study follows SHRP 2 Reliability Project L04 to assign traffic incidents/work zones independently based on the vehicle miles traveled (VMT) under a specific network variation combination. In other words, the larger the VMT of road  $l$ , the higher the probability  $p(l)$  that an incident/work zone will occur. The VMT of road  $l$  is the product of its length ( $\text{len}_l$ ) and the annual average daily traffic ( $\text{AADT}_l$ ). Monte Carlo simulation is conducted to generate road samples based on the discrete probability distribution until the incidents/work zones reach their occurrence frequency for a day in a network scenario.

$$p(l) = \frac{\text{len}_l \times \text{AADT}_l}{\sum_l \text{len}_l \times \text{AADT}_l} \quad (4.3)$$

Finally, the probability of a network scenario  $k$  built by the previous steps,  $p(\text{TN}_{1,2,\dots,i,j})_k$ , should be normalized to satisfy  $\sum_k p(\text{TN}_{1,2,\dots,i,j})_k = 1$ , which applies finite traffic network scenario simulations to represent the uncertainty of scenarios, as shown in Equation 4.4.  $\{1, 2, \dots, n\}$  is the scenario set.

$$p(\text{TN}_{1,2,\dots,i,j})_k = \frac{p(\text{TN}_{1,2,\dots,i,j})_k}{\sum_k p(\text{TN}_{1,2,\dots,i,j})_k}, k \in \{1, 2, \dots, n\} \quad (4.4)$$

### 4.2.3 Charging Profile Estimation

This section first provides the definition of future electrification levels and initial SOC, then explains how a single charging session is defined using explicit charging rules.

#### Future Electrification Levels

Future electrification levels estimate the potential transportation electrification level, considering the use of renewable energy and grid system updates. In this study, we use two electrification levels: moderate and high. Each level represents different EV market shares and the proportion of residential and workplace chargers. The electrification levels used in this study are based on the National Renewable Energy Laboratory (NREL) study [50].

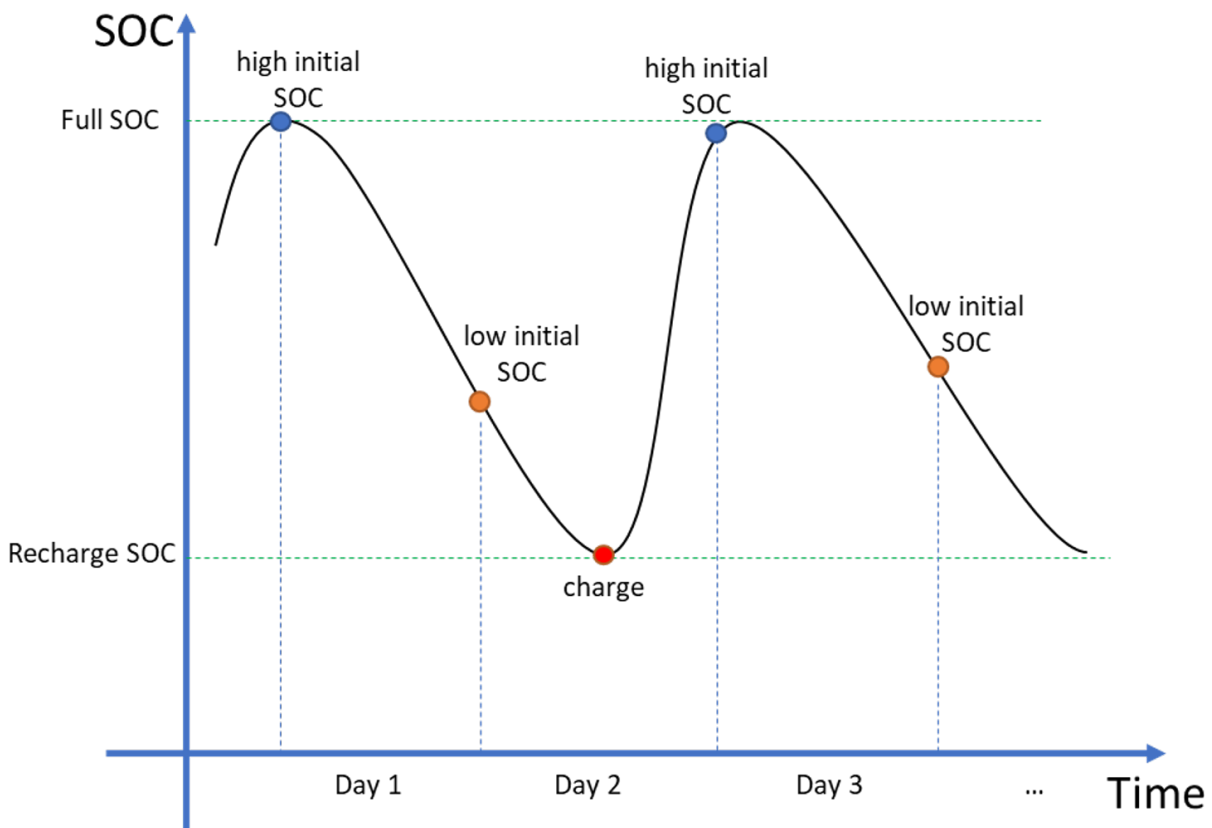


Figure 4.2: Initial SOC definition



## Initial SOC

Due to the variation in EV users' charging frequency, the initial SOC of an EV varies on a given day [246]. Therefore, we can expect a large deviation in charging needs throughout the day by selecting different initial SOC distributions across EV users. For instance, when most EV users start the day with a high SOC, the public charging needs could be very low, as the remaining range is sufficient to cover the whole day's trips. In contrast, when most EV users have a low SOC at the beginning of the day, the potential need for public charging could be considerably high. According to [246], the average charging frequencies for BEV and PHEV are 1.5 and 1.2 days, respectively. This means an EV would recharge about a day and a half after the last charge, as illustrated in Fig. 4.2. In other words, we can define two initial SOC states for EV users: (1) high initial SOC, or fully charged status (day one of a charging cycle); and (2) low initial SOC, or partially charged status (day two of a charging cycle). The high initial SOC scenarios can be regarded as the lower bound condition for charging demand since EVs start the day with a fully charged battery and are less likely to need public chargers. Meanwhile, the low initial SOC scenarios can be considered the upper bound condition for charging demand, as EVs start the day with partially charged states and are more likely to need public charging services.

## Charging Rule Definition

The charging rule defined in this study synthesizes literature associated with real-world charging behaviors [246], private/public charger distribution and disparities [21], macroscopical charging infrastructure assessment and projection [271], and EV battery performance [64]. Fig. 4.3 illustrates the logical decision-making process for applying destination-based public charging events. The core idea is to fully utilize the dwell time of scheduled activities to conduct charging events without interrupting the original schedule, unless the required charging time is insufficient for adjacent trips.

Charging decisions are made at the destination of each trip to determine whether to start a charging session at the destination. All charging rules can be categorized into two groups: (1) compulsory charging and (2) conditional charging.

- **Compulsory charging** ensures that all BEVs can complete all trips throughout the day by preventing BEVs from remaining in a very low SOC state.
- **Conditional charging** is implemented for both BEV and PHEV users when the external conditions are feasible to carry out a charging event without an urgent need for battery recharge as in compulsory charging. While most charging decisions are made at a trip destination, for trips that are the first trips of the tour and are BEV trips, an additional decision is made at the origin to ensure the EV's SOC is sufficient for the first trip.

The compulsory and conditional charging rules work simultaneously to ensure EVs always have enough electricity to travel and avoid impractical or excessive charging behaviors. The following criteria define charging behaviors:

- A compulsory charging decision is made when either of the following criteria is met:
  1. The SOC is lower than the minimum acceptable SOC.
  2. The SOC is insufficient for the energy required for the next adjacent trip.
- A conditional charging decision is made when all of the following criteria are met simultaneously:
  1. The SOC is below the anxiety SOC.
  2. The dwell time is long enough to recharge the EV to 80% SOC (EV maximum health SOC).
  3. The charging fee is cheaper than that at the next destination.

- When the EV is in a compulsory charging condition and the dwell time is insufficient for the required charging, a minimum charging period is added to the current destination's dwell time to fulfill the charging electricity requirement. All following trip schedules for this agent would be delayed accordingly.

Note that some workers in the synthetic population are identified as workplace charger users. The charging events carried out by these EV users at their workplace will still be simulated but will not be counted as charging demand for public chargers.

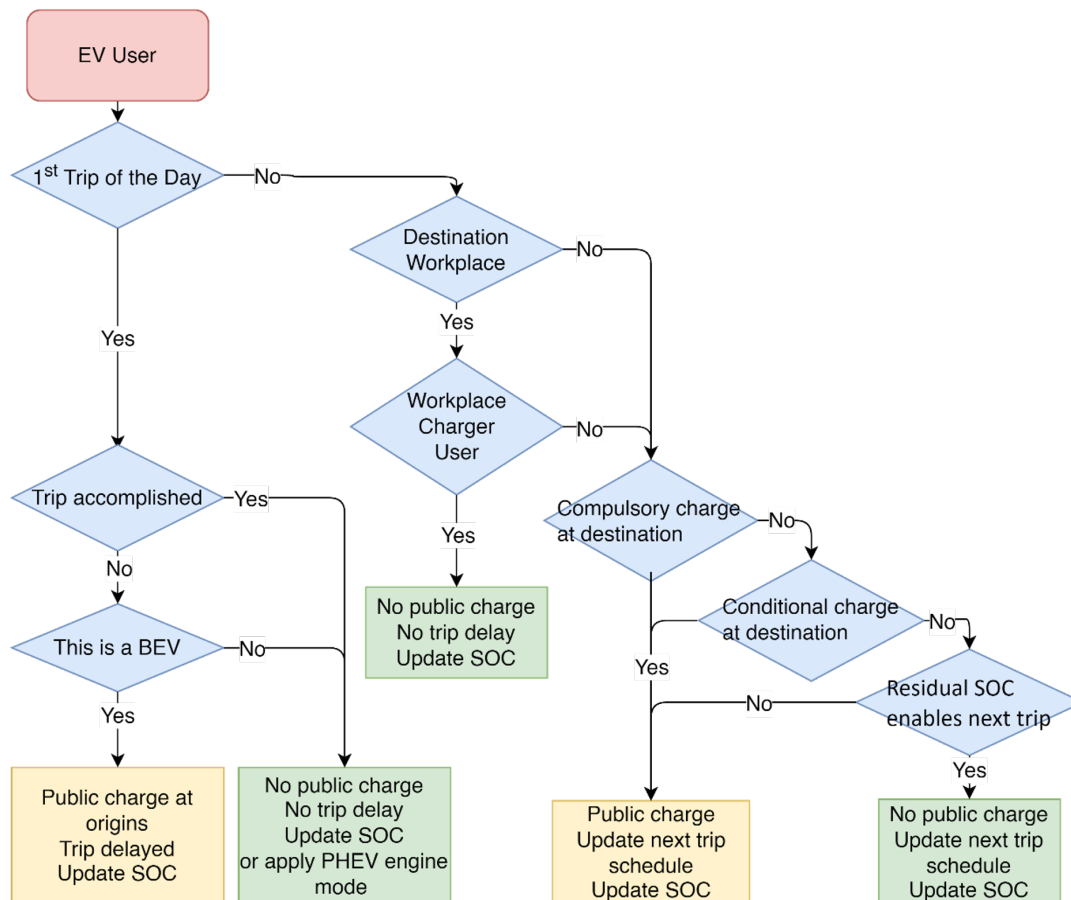


Figure 4.3: The public charging decision-making process

To sum up, the charging demand profile that comprises every single charging need across the day is generated by applying the charging rules to each travel profile provided in Section

4.2.2 under various combinations of electrification levels and initial SOC. Based on the individual-level charging demand, the temporal and spatial estimation of total charging needs and charging energy load at a higher level, such as TAZ or CT, can be accomplished.

## 4.2.4 Demand Aggregation and Equity Assessment

### Charging Demand Aggregation and Charger Estimation

Sections 4.2 through 4.3 provide a complete process of calculating the EV charging demand under combinations of different electrification levels, initial SOC, daily travel demand, and traffic network condition variations. Furthermore, from a planning perspective, it is more important to provide an overall estimation of charging demand. For instance, given a certain electrification level, what is the upper and lower bound estimation of charging demand for each zone and for the whole region? Or what is the estimated number of chargers required to satisfy a certain level of upper bound charging demand? This section (Section 4.2.4) will primarily discuss these questions.

#### *(1) Demand aggregation under each electrification level*

The first step in demand aggregation is to determine what fraction of simulated scenarios should be used for generating aggregated demand. The selection of simulated scenarios depends on the occurrence probabilities of the scenarios, as defined in Equation 4.6. Our goal is to avoid introducing bias in the ultimate estimation from scenarios with very low occurrence probabilities. We define an index called acceptance level (AL) to eliminate the impact of scenarios with minimal occurrence frequency, as they are not common cases that need serious attention from planners. The idea of AL is similar to the concept of risk acceptance in risk management, where a business or individual acknowledges that the potential loss from a risk is not significant enough to warrant effort to avoid it. AL is implemented by ensuring that the sum of occurrence probabilities of those high-frequency scenarios is at or close to AL, as shown in Equations 4.5 and 4.6.  $TN'_{(1,2',\dots,n')}$  in Equation 4.5 is the network scenario set

with the descending order of occurrence probability  $p(\text{TN}_{(1,2,\dots,i,j)})_k$ , reindexed by  $1', 2', \dots, n'$ . Equation 4.6 requires that the public charging demand profile prediction satisfies the top  $1'$  to  $m$  scenarios in  $\text{TN}'$ , which occupies frequent network scenarios up to AL. AL is adjustable depending on the number of simulated scenarios, evaluated case by case.

$$\text{TN}'_{(1',2',\dots,n')} = \text{TN}_{(1,2,\dots,n)} \left[ \text{sorted}_{\text{desc}} \left( p(\text{TN}_{(1,2,\dots,i,j)})_k \right) \right] \quad (4.5)$$

$$\text{AL} \approx \sum \text{TN}'_{1',2',\dots,m}, \quad m \leq n \quad (4.6)$$

After applying AL, the predicted charging demand for a zone under an electrification level  $e$ ,  $D_{\text{zone}}^e$ , is obtained as the maximum of the charger numbers for this zone across scenarios  $1'$  to  $m$ , as shown in Equation 4.7.

$$D_{\text{ct}}^e = \max \left( \text{TN}'_{(1',2',\dots,m)} \left\{ D_{\text{zone}}^{(1',e)}, D_{\text{zone}}^{(2',e)}, \dots, D_{\text{zone}}^{(m,e)} \right\} \right) \quad (4.7)$$

## (2) Charger number estimation

Based on the charging demand estimation, this study proposes an approach to estimate the minimum number of chargers required to satisfy the charging demand of each zone. By assuming that a charger can serve multiple charging needs as long as they do not overlap in time, the charger number estimation problem is transformed into an equivalent problem of finding the maximum overlap of charging sessions for each zone over 24 hours. Fig. 4.4 provides an example of the relationship between the number of charging demands and the number of required chargers. As seen in Fig. 4.4, there are five charging needs scheduled between time  $t_1$  and  $t_2$ , and each charging session is represented as a horizontal bar with an exclusive start and end time. The maximum overlap between the five charging sessions during time  $t_1$  to  $t_2$  is three, meaning that three chargers are sufficient to serve the five charging needs without a waiting time between two adjacent charging sessions. Similarly, the minimum

number of chargers required to serve all charging demands for a day in a given zone can be derived by calculating the maximum overlap between each charging session over 24 hours.

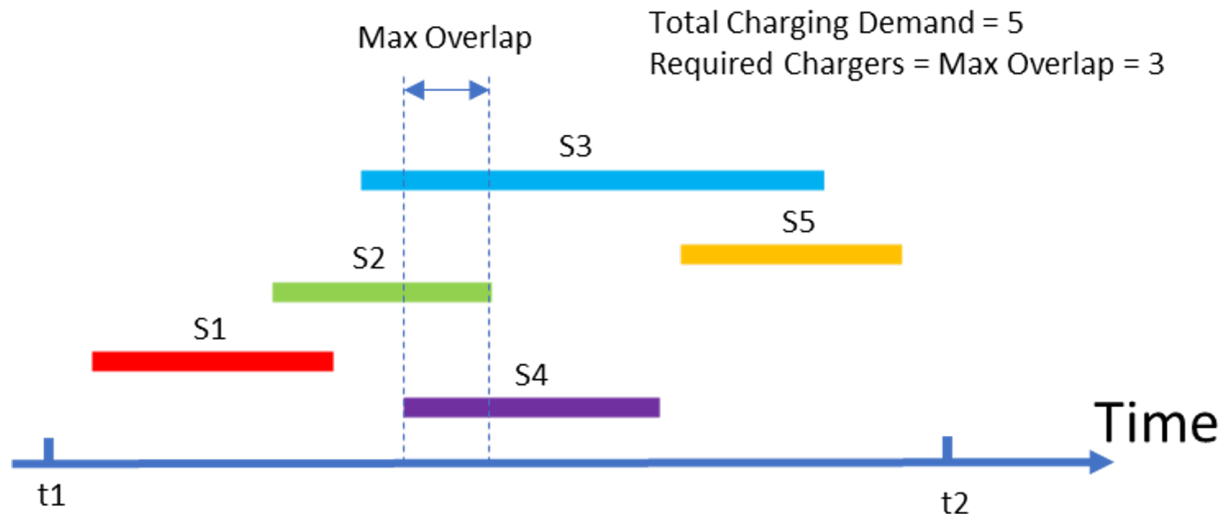


Figure 4.4: Relation between charging demand and required chargers

Considering the difference in charging needs between high and low initial SOC, the charging needs per zone may vary across days, which might be much lower than the predicted maximum charger number. To account for the utility rate of chargers and installation budgets, we define a satisfaction level (SL) to represent to what extent the predicted number of chargers can be satisfied between the lower and upper bounds of charge demands. For a given electrification level  $e$ ,  $SL = 100\%$  means the installation of chargers is not restricted by the budget and satisfies the upper bound of estimated chargers for each zone in the system.  $SL = 0\%$  means the investment can satisfy the lower bound of estimated chargers in the entire network. Other values in-between (0% and 100%) represent the tradeoff between improving system charging service and managing installation costs. The budget constraint is formulated as Equation 4.8.

For a given electrification level  $e$ , the required number of chargers for each zone is in the range of  $[C_{\text{zone}}^{(e,\text{low})}, C_{\text{zone}}^{(e,\text{up})}]$ , where  $C_{\text{zone}}^{(e,\text{low})}$  and  $C_{\text{zone}}^{(e,\text{up})}$  refer to the lower and upper bounds

of the estimated number of chargers ( $C_{ct}^e$ ), considering the maximum overlap of charging sessions in electrification level  $e$ .  $U_{ins}$  is the unit installation cost of a charger.  $I_{bud,e}$  is the installation investment budget for electrification level  $e$ .

$$\sum_{zone} U_{ins} \times [SL \times (C_{zone}^{(e,up)} - C_{zone}^{(e,low)}) + C_{zone}^{(e,low)}] \approx I_{bud,e} \quad (4.8)$$

Equations 4.9 and 4.10 present the final projected zone-level and network-level public charger numbers ( $T_{zone,e}$  and  $T_{net,e}$ ) of the transportation system under a given electrification level  $e$  to maximize system charging service under financial restrictions on installation.

$$T_{zone,e} = SL \times (C_{zone}^{(e,up)} - C_{zone}^{(e,low)}) + C_{zone}^{(e,low)} \quad (4.9)$$

$$T_{net,e} = \sum_{zone} T_{zone,e} \quad (4.10)$$

## Charging Equity Measurements

In order to assess the equity performance of public charging accessibility at the projected electrification level, this subsection evaluates social equity in accessibility across various groups of zones for the target year and compares the performance with the current year.

### (1) Socio-demographic classification

We selected a couple of representative socio-demographic variables to divide the whole study area into different groups of zones. These variables are also adopted by similar EV charging equity studies. The following equity analysis will be conducted based on these different socio-demographic groups. The definition of the socio-demographic variables can be found in Table 4.1.

### (2) Charging equity metrics

The assessment adopts two accessibility metrics to indicate the level of accessibility to public charging across the different socio-demographic groups.

Table 4.1: The definition of the socio-demographic variables

Variable Name	Definition	Grouping Examples
Disadvantaged Communities	Areas that most suffer from a combination of economic, health, and environmental burdens.	1. Disadvantaged communities (DAC) 2. Non-Disadvantaged communities (non-DAC)
Zonal Income	The median household income of the zone.	1. Low income 2. Medium low income 3. Medium high income 4. High income
Zonal Major Ethnicity	The ethnicity with greater than 50% of the population in the zone.	1. White 2. Black/Hispanic 3. Asian 4. No majority

**Charger access probability (CAP).** For a single CT, if it has at least one public charger within its boundary, it is defined as having public charging access. For a group of CTs, CAP is calculated as the portion of CTs with at least one charger among all tracts given a selected grouping of socio-demographic features.

$$CAP_i = \frac{n_i}{N_i} \quad (4.11)$$

where  $i$  refers to the  $i$ th socio-demographic group,  $n_i$  represents the number of CTs with at least one public charger in group  $i$ , and  $N_i$  is the number of CTs in group  $i$ .

In addition, we use the Gini index [253] to indicate the quantitative disparity in charging accessibility across different groups with regard to CAP.

**Gini index across socio-demographic groups.** The Gini index is the most frequently used equity indicator for the distribution of accessibility and other effects. For a given socio-demographic metric in a selected year, a Gini index is calculated to indicate the level of disparity across socio-demographic groups. In general, the greater the Gini index, the larger the disparity. The Gini index is calculated for both 2022 and 2035 to compare their charging equity performance.

$$\text{Gini Index} = \frac{1}{2n^2 \bar{X}} \sum_{i=1}^n \sum_{j=1}^n |X_i - X_j| \quad (4.12)$$



where  $n$  refers to the number of socio-demographic groups (e.g., income groups or ethnicity groups),  $X_i$  refers to the value of the selected equity performance metric for socio-demographic group  $i$ , and  $\bar{X}$  is the mean of all  $X_i$ .

(3) *Control variable*

To visualize the public charging disparities in depth, the charging equity metrics for every socio-demographic group are displayed by varying the value of the control variable. The multi-family housing unit (MFHU) rate is adopted as the control variable because higher MFHU concentrations in a CT lead to a greater public charger need due to lower access to dedicated parking and home chargers, which has been discussed in other public charging studies. The MFHU rate is calculated as the total number of MFHU units divided by the total number of housing units per CT. To investigate the charging equity performance at different MFHU levels, the MFHU rate is divided into four groups with identical intervals:  $[0, 0.25]$ ,  $(0.25, 0.5]$ ,  $(0.5, 0.75]$ , and  $(0.75, 1]$ .

$$\text{MFHU Rate}_i = \frac{N_{\text{mfhu},i}}{N_i} \quad (4.13)$$

where  $N_{\text{mfhu},i}$  refers to the number of MFHU units in the  $i$ th CT, and  $N_i$  refers to the total number of units in the  $i$ th CT.

To summarize, Sections 4.2.1 through 4.2.4 provide a complete description of the framework for public charging demand estimation that accounts for different future electrification levels, travel demand and traffic condition variations, and charging behavior heterogeneity. As an expansion of Fig. 4.1, a more specific illustration of the major steps involved in the charging demand prediction system developed in this study is shown in Fig. 4.5.

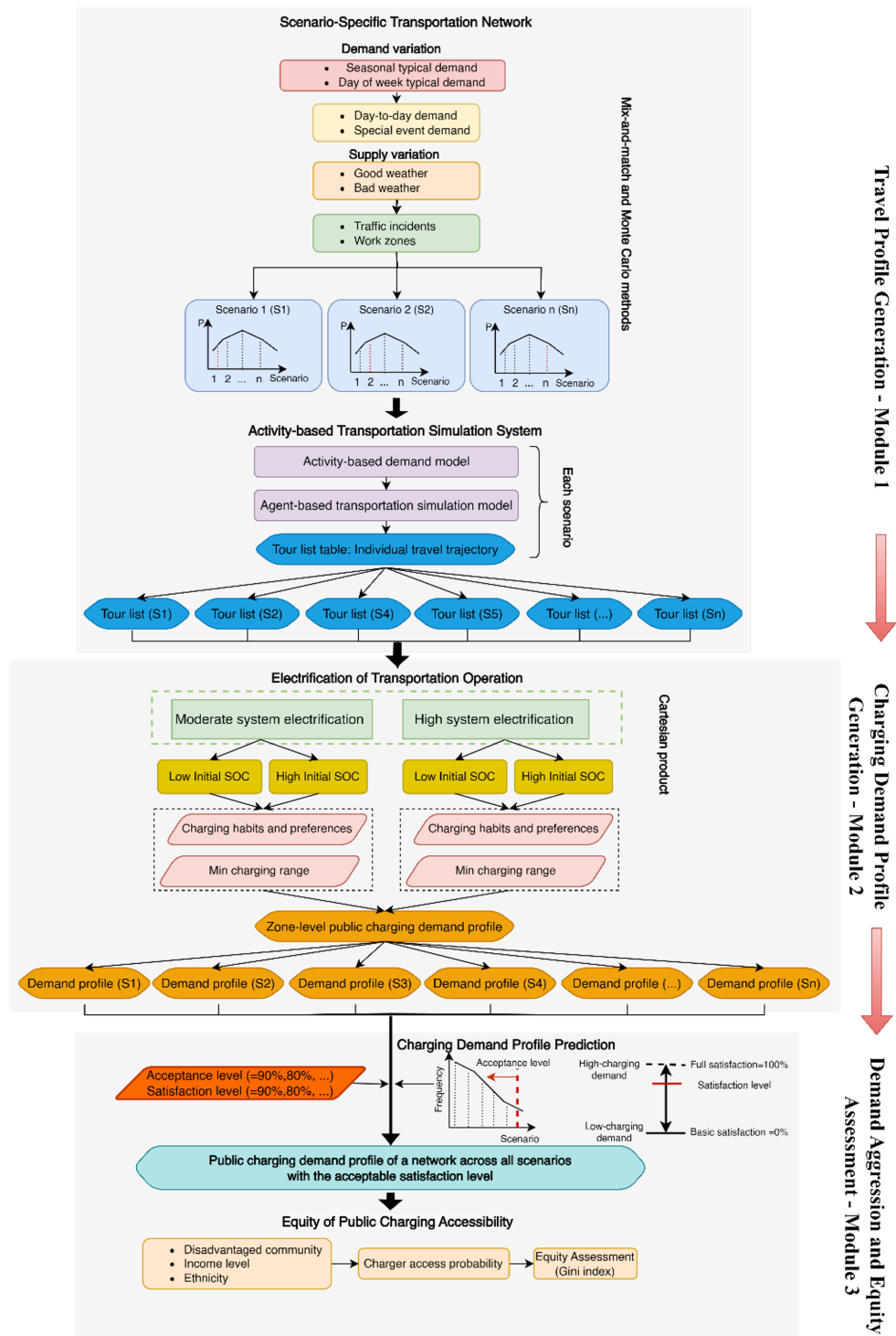


Figure 4.5: Major Steps in the Charging Demand Prediction Model

## 4.3 Case Study and Data Specification

The case study focuses on the large-scale transportation system of Los Angeles (LA) County in California. The studied transportation system contains 2,342 CTs and 38% of them are DACs, with all primary and secondary roads, a population of over 10 million, and more than 6 million vehicles, with only 1.23% EV market share in 2022, see Fig. 4.6.

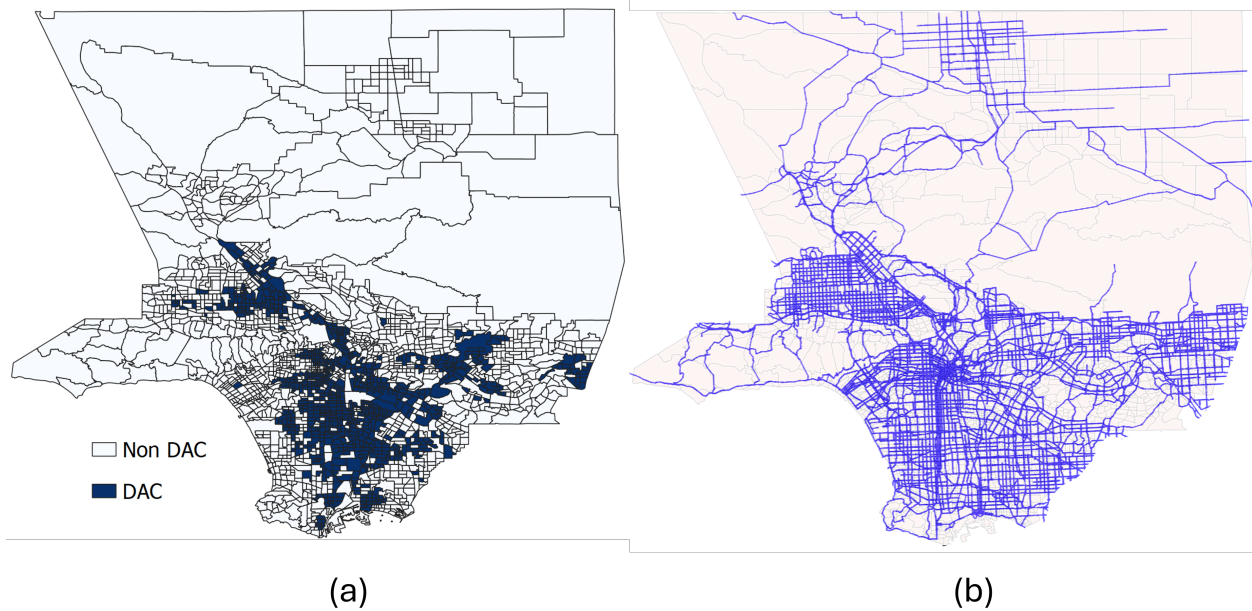


Figure 4.6: (a) Census tracts of LA county (CTs); (b) The road transportation network

### 4.3.1 Travel Data Specification

This section follows Section 4.2.2 to simulate transportation network scenarios in 2035 via historical traffic data. Given the data availability, the case study practically concerns the scenario components of the day of the week and traffic incidents (i.e., traffic interruptions and road closures) from both traffic demand and network supply sides. Fig. 4.7 provides an illustration of the travel demand distribution across the 24 hours of weekday and weekend scenarios. As shown in Fig. 4.7(a), the weekday demand has two notable peaks: the AM peak

near 8 AM and the PM peak near 4 PM, while the weekend demand shows no significant peak. Fig. 4.7(b) and Fig. 4.7(c) show a breakdown of travel demand by trip purpose. It can be seen that the number of work trips drops significantly on weekends compared to weekdays. On the other hand, the non-work trips on weekdays and weekends do not show a significant difference in magnitude, though their temporal patterns differ.

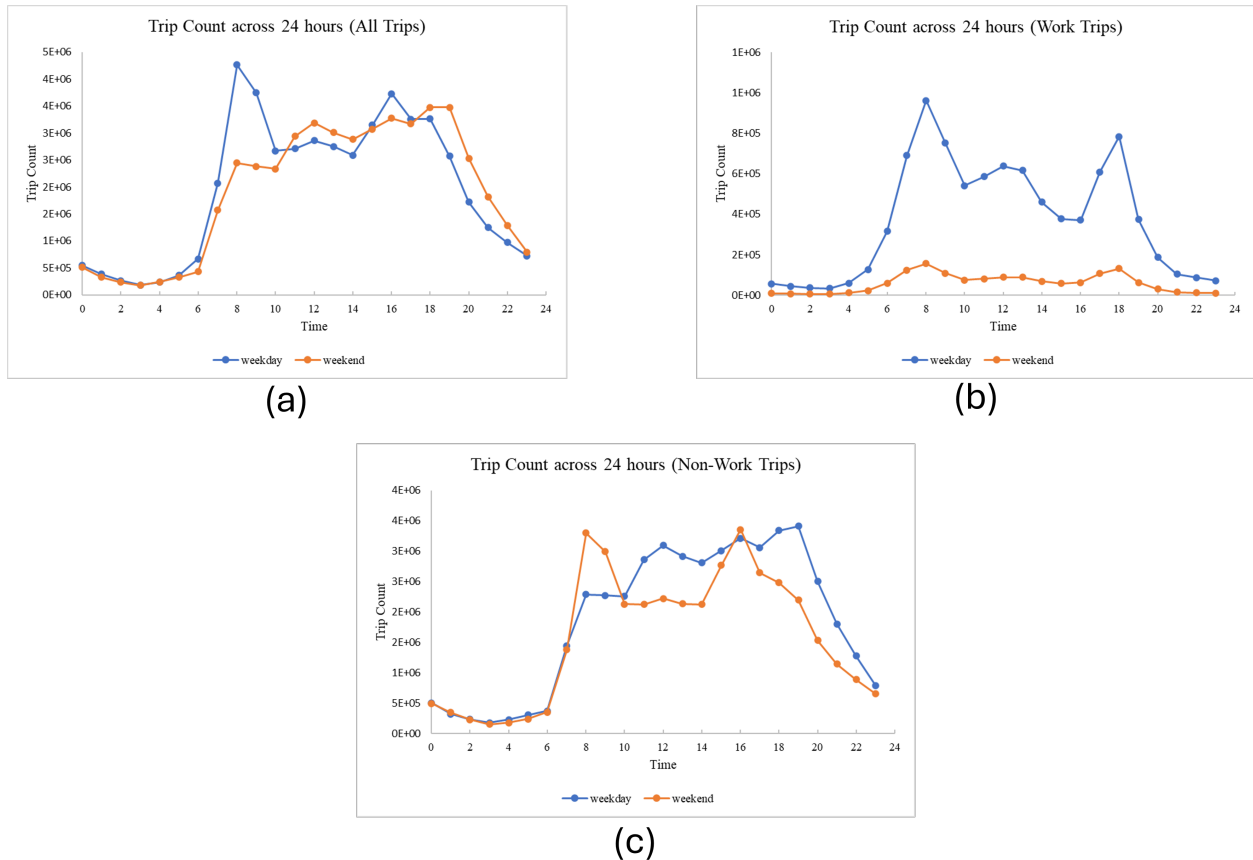


Figure 4.7: Travel demand temporal distribution by trip purposes: a) All trips, b) Work trips, and c) Non-work trips

Generally, the data parameters and distributions associated with the transportation network scenario generation in the case study are listed in Table 4.2. Then, the mix-and-match approach was used to merge the weekday/weekend and weather status, and the Monte Carlo approach was conducted based on the distributions in Table 4.2 to produce the incident frequency of a day under each group of the day of the week, weather status, start time,

duration, and related road capacity loss of each traffic incident. This study assumes that weather and people’s traffic behavior will remain consistent over the next decade. However, due to the significant growth rate (15.2%) of the population in 2035 estimated by SCAG, the daily incident number in 2035 is linearly projected by the product of the incident frequency in 2022 and the 2035 population growth rate. Finally, the case study simulates 10 network scenarios (from scenario S1 to S10) of the LA County transportation system in 2035, as shown in Table 4.3. The 10 scenarios are assigned based on the probability of each condition group in Table 4.2, with higher probability groups having more scenarios; for example, 6 out of 10 scenarios are in the weekday with no-rain/snow group.

### **4.3.2 Electrification and Charging Data Specification**

To ensure a valid comparison of results among network scenarios and transportation electrification operations, EV-related assumptions and parameters for the case study are defined in the following bullets and in Table 4.4. Table 4.4 specifically lists the parameter values and literature associated with EV batteries, charge prices, and charger efficiency.

#### **Electrification Data**

The EV market share rate for each CT in LA County in 2035 is assumed to have the same distribution across all CTs as in 2022, with a linear projection of the total number of EVs from 2022 to 2035. The projected number of EVs is 1.16 million under the 2035 moderate electrification level and 3.09 million under the 2035 high electrification level. The spatial distribution of the 2035 EV adoption rate is shown in Fig. 4.8.

Referring to California electrification research, under a moderate electrification level [14], 80% of EV residents are assumed to have home chargers; under a high electrification level, potentially 70% of EV residents have home chargers. The total number of home chargers is first divided into four batches by residency type. Then, within each residency type, the

Table 4.2: Parameters and distributions of LA county transportation network scenarios

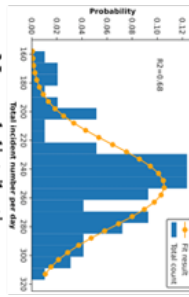
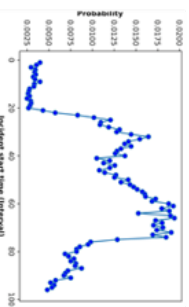
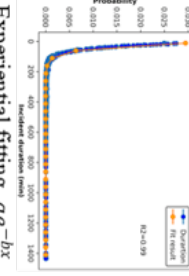

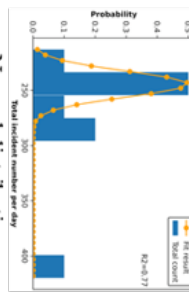
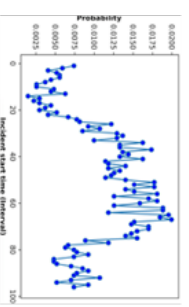
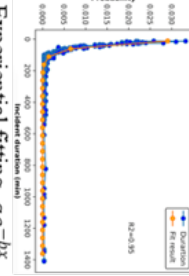
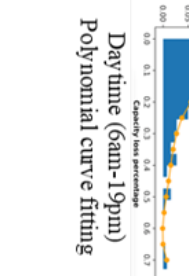
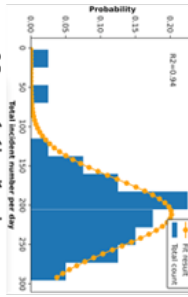
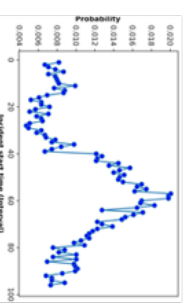
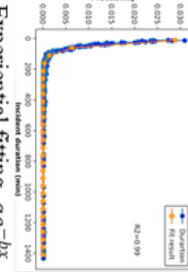
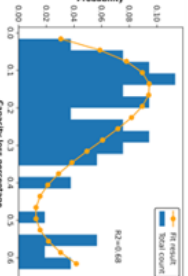
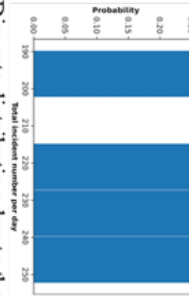
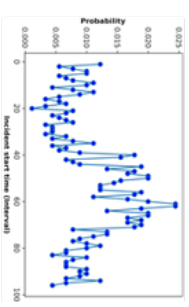
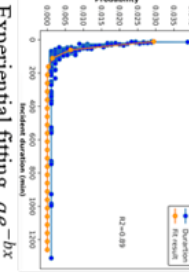

Condition group for incidents (Probability)	Daily incident frequency	Incident start time (15mins time interval)	Incident duration (mins)	Capacity loss percentage
<p>Weekday, no rain/snow (<math>P=0.64</math>)</p>	 <p>Normal distribution <math>X \sim N(252, 29)</math></p>	 <p>Discrete distribution</p>	 <p>Exponential fitting, <math>ae^{-bx}</math> <math>a = 0.0397, b = 0.0297</math></p>	 <p>Daytime (6am-19pm) Polynomial curve fitting</p>
<p>Weekday, rain/snow (<math>P=0.07</math>)</p>	 <p>Normal distribution <math>X \sim N(244, 12)</math></p>	 <p>Discrete distribution</p>	 <p>Exponential fitting, <math>ae^{-bx}</math> <math>a = 0.0373, b = 0.0283</math></p>	 <p>Nighttime (19pm-6am) Polynomial curve fitting</p>
<p>Weekend, no rain/snow (<math>P=0.27</math>)</p>	 <p>Normal distribution <math>X \sim N(211, 44)</math></p>	 <p>Discrete distribution</p>	 <p>Exponential fitting, <math>ae^{-bx}</math> <math>a = 0.0398, b = 0.0301</math></p>	 <p>Nighttime (19pm-6am) Polynomial curve fitting</p>
<p>Weekend, rain/snow (<math>P=0.03</math>)</p>	 <p>Discrete distribution due to the few cases</p>	 <p>Discrete distribution</p>	 <p>Exponential fitting, <math>ae^{-bx}</math> <math>a = 0.0382, b = 0.0289</math></p>	 <p>Nighttime (19pm-6am) Polynomial curve fitting</p>

Table 4.3: The simulated transportation network scenarios in 2035 with their incident numbers and normalized occurrence probability

Weekday/weekend and weather	Scenario ID	Daily incident number	Normalized P
Weekday, no-rain/snow	S1	288	0.125
	S2	311	0.102
	S3	297	0.122
	S4	288	0.125
	S5	322	0.079
	S6	317	0.090
Weekday, rain/snow	S7	282	0.066
Weekend, no-rain/snow	S8	242	0.142
	S9	215	0.123
Weekend, rain/snow	S10	274	0.026

Table 4.4: EV-related parameters for the 2035 electrified transportation system in LA county

Input	Unit	Value	EV type	Source
Battery size	kWh	$X \sim Uni(73,163)$	BEV	Crisostomo et al. 2020
		$X \sim Uni(11,18)$	PHEV	
Minimum SOC tolerance range	%	$X \sim Uni(10\%, 20\%)$	BEV	Wood et al. 2018
		$X \sim Uni(30\%, 40\%)$	PHEV	Mahmoodi-k 2021
Anxiety SOC range	%	$X \sim Uni(50\%, 60\%)$	EV	Tal et al. 2020
Plug-off SOC	%	80% for short activity durations 100% for long activity durations	EV	Tal et al. 2020
Charge price	\$/kWh	12am-8am: 0.32-0.44(\$/kWh) 8am-4pm: 0.4-0.55(\$/kWh) 4pm-9pm: 0.48-0.66(\$/kWh) 9pm-12am: 0.4-0.55(\$/kWh)	EV	<a href="#">EVgo Services LLC</a> . 2022
Pedestal mounted L2 charger average installation cost	\$	\$1450, one plug per charger	EV	Smith et al. 2015
L2 charger charging level/speed	kW	6.6kW	EV	Smith et al. 2015

home chargers are randomly assigned to households. The share of home chargers in each residency type is determined by the California Energy Commission’s report on future PEV infrastructure projections [271]. As estimated by NREL [50], 14% (moderate electrification) or 28% (high electrification) of EV residents in 2035 can access workplace chargers during their work activities.

The focus of this study is on Level 2 (L2) chargers due to their wide adoption in the public charger market (nearly 90%) and their wide installation with less technical requirements on the grid system. This study assumes that one charger has one plug for users.

### **Charging Data**

Since 80% of full SOC is the cap of the healthy operational state for prolonging the life of EVs and reducing battery degradation, for EV (BEV and PHEV) residents with home charger access, we assume they always start their first trips of the day with 80% SOC, regardless of whether it is a high- or low-initial SOC scenario.

For EV residents without access to home chargers, due to limited charging opportunities (only public and workplace charging), two SOC values are designed for their first trips of the day in the case study. In high initial SOC scenarios, it is assumed that all EVs are charged to a high SOC before their first trips, and that 80% of full SOC (healthy battery state) is assigned to EVs before the first travel on average. In low initial SOC scenarios, the second day after charging leads to a low SOC for the start of daily trips. In this condition, on average, EV residents without home chargers start their day with 40% of full SOC (half of the first day’s SOC) to reflect the first day’s energy consumption.

The EV energy consumption rate (ECR) is typically determined by an EV traveling a certain distance and the corresponding travel characteristics, such as travel speed, terrain conditions, temperature, waiting time, etc. However, for a large-scale road network with tens of thousands of EVs, a more general assumption is often made to use the average travel speed to indicate the ECR (e.g., kWh/km) for a specific EV trip. This study adopted the empirical



equation for the relationship between average speed and ECR derived by [81] to calculate the electricity consumption of each EV trip. The equation is based on data collected from GPS loggers installed in 200 BEVs used by 741 drivers for 276,102 trips covering about 2.3 million km traveled. The equation is shown in Equation 4.14, where  $v$  refers to the average travel speed. The relationship between the average travel speed and the ECR indicates that driving at very slow or very fast speeds increases the ECR, and the most energy-efficient driving speed is between 45 and 56 km/h.

$$\text{ECR} = -0.0002v^3 + 0.071v^2 - 5.6558v + 308.82 \quad (4.14)$$

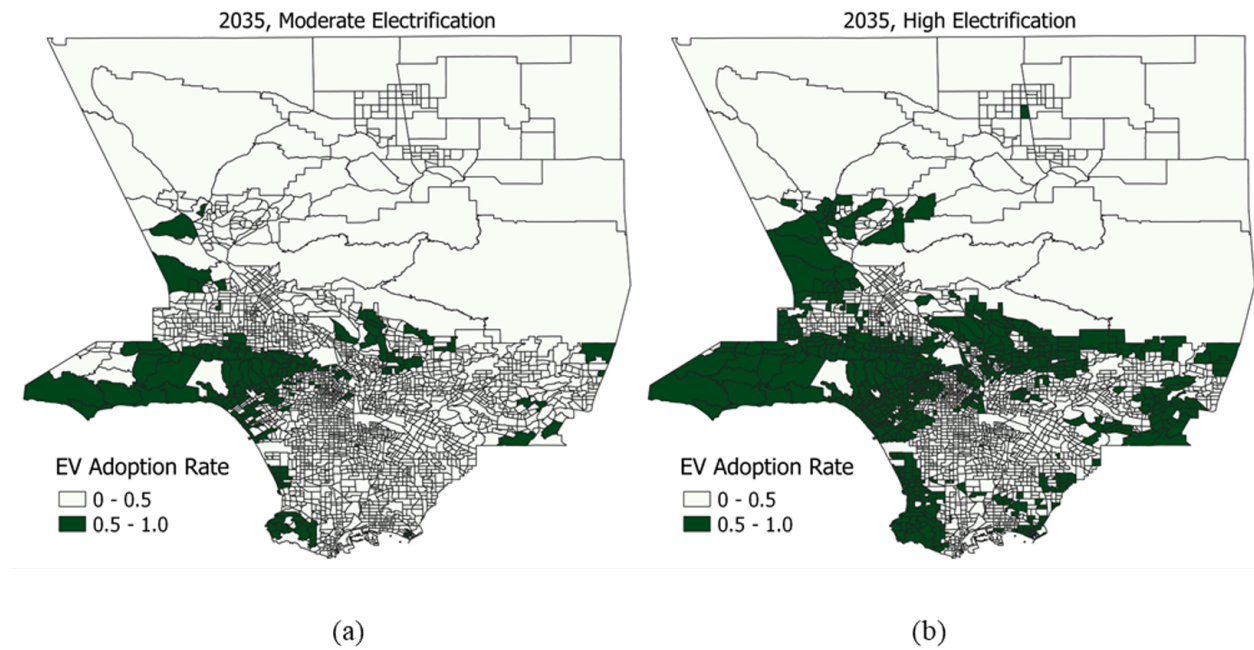


Figure 4.8: 2035 EV adoption rate across CTs under (a) moderate electrification level and (b) high electrification level

## 4.4 Results and Discussion

### 4.4.1 Network-Level Charging Demand Analysis

Table 4.5 provides the network-level outcome associated with the CREATE application to predict the needs for public charging (specifically L2 chargers) in the electrified transportation network of LA County in 2035. This table presents the mean values and standard deviations to illustrate the prediction variation influenced by different scenarios. The prediction results are divided by weekday and weekend, given their distinctive differences in travel demands. Table 4.5 and most of the following discussions focus on the charging demand during the peak demand hour, i.e., the one-hour period with the maximum charging demand throughout the day.

Table 4.5: Network-level L2 public charging demand for LA county in 2035 under all scenarios

Electrification level	Initial SOC	Peak charging demand (count)		Peak energy consumption (kWh)		Average duration (hour)	
		Weekday scenarios (mean (std))	Weekend Scenarios (mean (std))	Weekday Scenarios (mean (std))	Weekend Scenarios (mean (std))	Weekday Scenarios (mean (std))	Weekend Scenarios (mean (std))
High	Low	305,460 (1,083)	95,913 (541)	1,642,508 (6,355)	503,355 (2,831)	2.73 (0.01)	2.80 (0.01)
	High	152,790 (820)	60,316 (244)	747,423 (4,000)	291,154 (1,092)	1.41 (0.01)	1.49 (0.01)
Moderate	Low	118,173 (970)	35,363 (440)	633,886 (4,101)	183,855 (2,623)	2.59 (0.01)	2.68 (0.02)
	High	65,207 (803)	23,300 (20)	321,981 (3,078)	112,563 (1,247)	1.44 (0.01)	1.53 (0.02)

The peak demand columns in Table 4.5 provide the range of charging demand during the peak demand hour for weekday or weekend scenarios. The charging demand under the high electrification level is about 2.3-2.6 times that under the moderate level, and under the same electrification level, the demand for the low SOC scenario is nearly 1.8-2.0 times that in the high SOC scenario. Additionally, the high values in the weekday peak demand column are primarily produced by relatively long-distance mandatory trips, whereas the low values in the weekend peak demand column are mainly due to the dominance of short-distance

non-mandatory trips.

The peak energy consumption represents the hourly electrification load when the peak charging demand appears. The energy consumption values in the weekday and weekend columns follow the same trend as the peak demand columns. The average duration in this table is the average charging time per charging event for an EV user under a scenario. The low initial SOC scenarios, due to the low SOC of EVs at the start of their daily travels, generate a longer charging duration for both high and moderate electrification levels on average. Another contributing factor is that the high electrification level has a higher EV market share, while the development of home chargers is constrained by land use, power restrictions, etc., leading more EVs to rely on public chargers, which results in longer public charging durations under high transportation electrification.

Although Table 4.5 does not show a significant network-level difference in public charging needs across scenarios for each column, the results in Sections 4.1.1 and 4.1.2 reveal that the temporal and spatial variations at the CT level are distinct.

### **Network-level Temporal Pattern Analysis**

Fig 4.9 demonstrates the temporal charging demand pattern of weekday and weekend scenarios under four electrification groups with a 1-hour time interval throughout the day. We calculated the hourly charging demand for all scenarios and used the average demand during each one-hour period across scenarios to represent the temporal charging pattern for both weekday and weekend scenarios. This process is repeated for four combinations of electrification levels and initial SOC.

As shown in Fig 4.9, under the same electrification levels, the weekday charging demand is significantly higher than the weekend demand across all time periods. The peak weekend demand is only about 25% of the peak weekday demand under the same electrification level. The peak of daily charging demand appears near 10 to 11 AM for weekday scenarios and 1 to 2 PM for weekend scenarios. The large gap between weekday and weekend charging

demand is due to the reduced work trips during weekends compared to weekdays. Since the number of weekday work trips is significantly larger than on weekends, and the average work trip distance is 68% longer than non-work trips, it is expected that EVs generate more charging needs during weekdays than weekends. The temporal charging energy load pattern is consistent with the temporal charging demand pattern. The peak of daily charging energy load appears around 10 AM for weekday scenarios and around 1 to 2 PM for weekend scenarios.

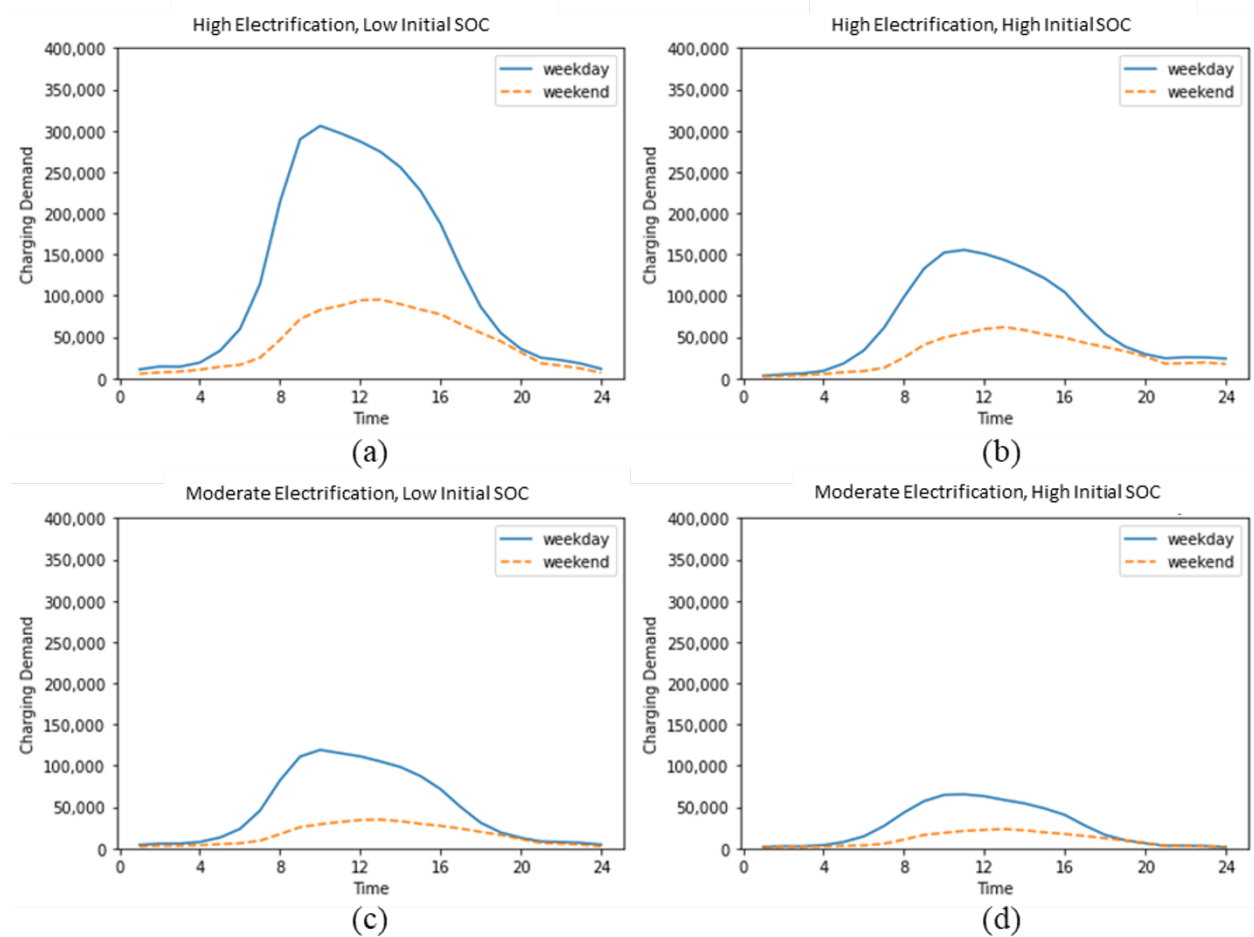


Figure 4.9: Temporal distribution of charging demands from weekday scenarios and weekend scenarios of (a) High electrification plus low initial SOC; (b) High electrification plus high initial SOC; (c) Moderate electrification plus low initial SOC; and (d) Moderate electrification plus high initial SOC

## Network-level Spatial Pattern Analysis

This subsection studies the spatial distribution of charging demands across electrification groups. Fig 4.10 shows the geographic regions and the spatial distribution of population density in LA County. As seen in Fig 4.10(b), the population is concentrated near central LA, San Fernando Valley, southern San Gabriel Valley, and the South Bay area. To illustrate the spatial distribution of charging demand more intuitively, we select the charging demand generated from the weekday scenarios as an example and calculate the average charging demand for each CT across different weekday scenarios to reflect the spatial pattern under four electrification groups.

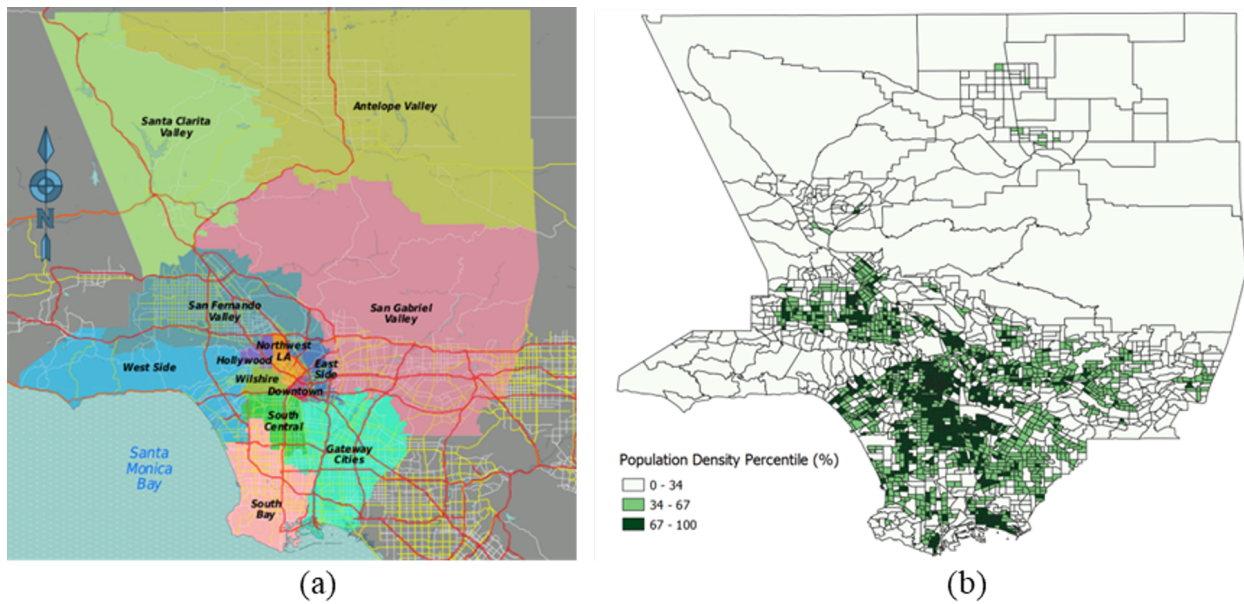


Figure 4.10: LA county (a) Geographic regions (Wikimedia, 2022); and (b) Population density distribution (California Air Resources Board (CARB), 2022)

The charging demand in Fig 4.11 is measured in area density, i.e., demands per square kilometer, for each CT. CTs are colored into two groups where dark red indicates high-density CTs and light red represents low-density CTs. The classification threshold is determined by the median charging needs density across CTs, i.e., 16 demands/1,000 inhabitants in the

weekday scenario of the high electrification, high charging-demand-day group. As shown in Fig 4.11, areas with higher charging demand density correspond to areas with higher population density in Fig 4.10(b). Among the four maps in Fig 4.11, Fig. 4.11(a) shows the most concentrated distribution of charging demand, where the high-density CTs occupy the majority of central and southern LA regions. In contrast, the demand density distribution is less concentrated under conditions with high initial SOC or a moderate electrification level. Notably, for the moderate electrification level, due to the low EV market share rate estimated, both high and low initial SOC scenarios exhibit lighter density distributions compared to the maps of the high electrification level.

#### 4.4.2 Zone-level Charging Demand Variation Across Scenarios

To analyze the demand variation of CTs across all test network scenarios, the mean and standard deviation of the peak charging demand across scenarios for each CT are calculated. As discussed in Table 4.4, the predicted charging demand differs under different electrification levels, days of the week, and daily charging demands. Thus, the sample scenarios applied in this subsection to emphasize zone-level demand variation under multiple scenarios are scenarios S1 to S7, weekday scenarios under high electrification level plus low initial SOC. We only demonstrate one combination of electrification levels and initial SOC because other combinations share similar characteristics. We then normalize the peak charging demand by dividing the peak demand by the total number of dwelled cars within the one-hour period of peak demand for each CT. Here, dwelled cars refer to cars that dwell in a zone within a one-hour period. The public charging demand distribution of the mean and the standard deviation for the sample scenarios is plotted in Fig 4.12.

As shown in Fig 4.12(a), the mean of average demand for weekday scenarios is 12.5 charging needs per 100 dwelled cars in a CT during the peak demand hour. In Fig 4.12(a), we can see that the standard deviations of the normalized charging demand are relatively



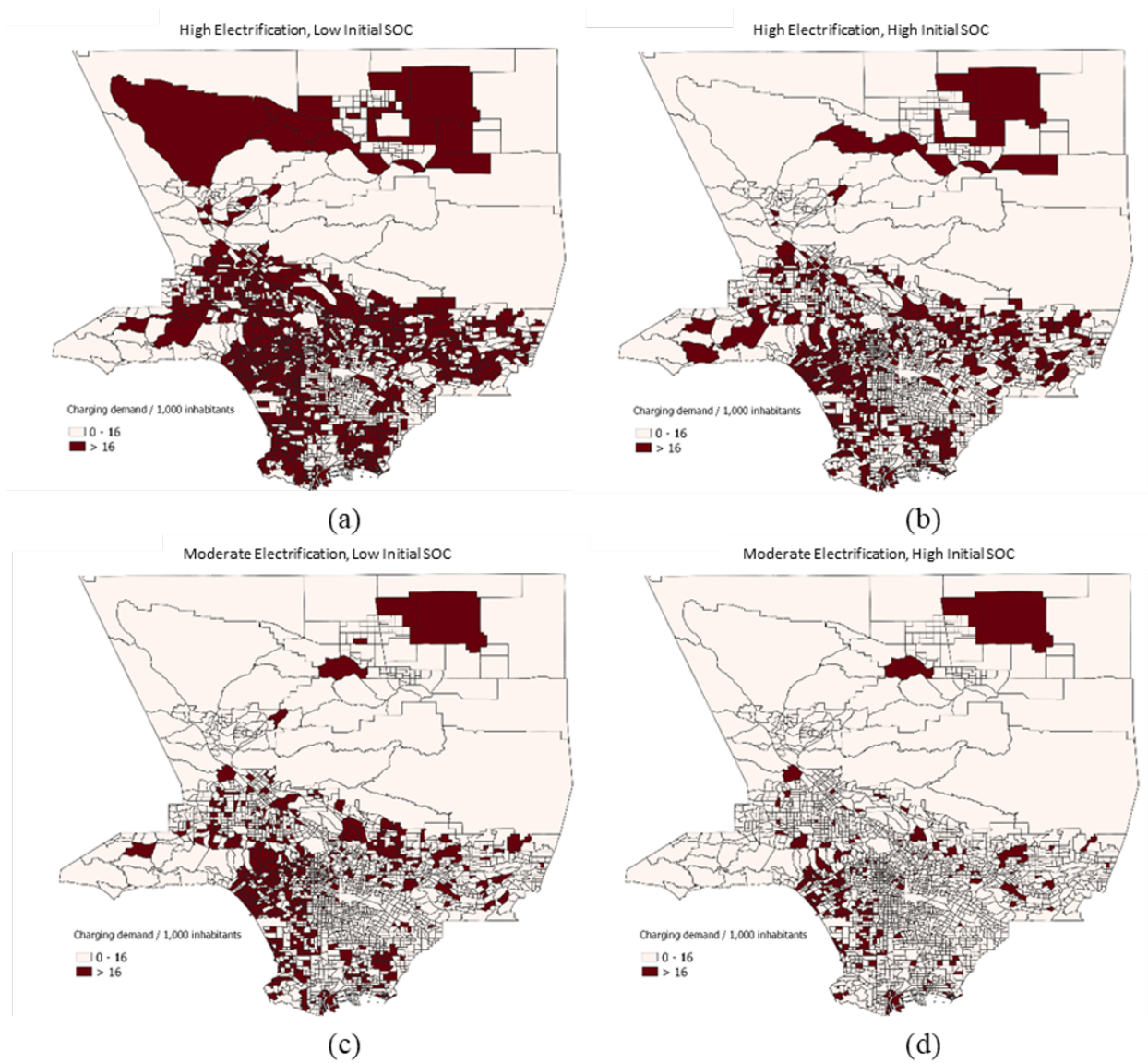


Figure 4.11: Spatial distribution of peak charging demand from weekday scenarios of (a) High electrification plus low initial SOC; (b) High electrification plus high initial SOC; (c) Moderate electrification plus low initial SOC; and (d) Moderate electrification plus high initial SOC

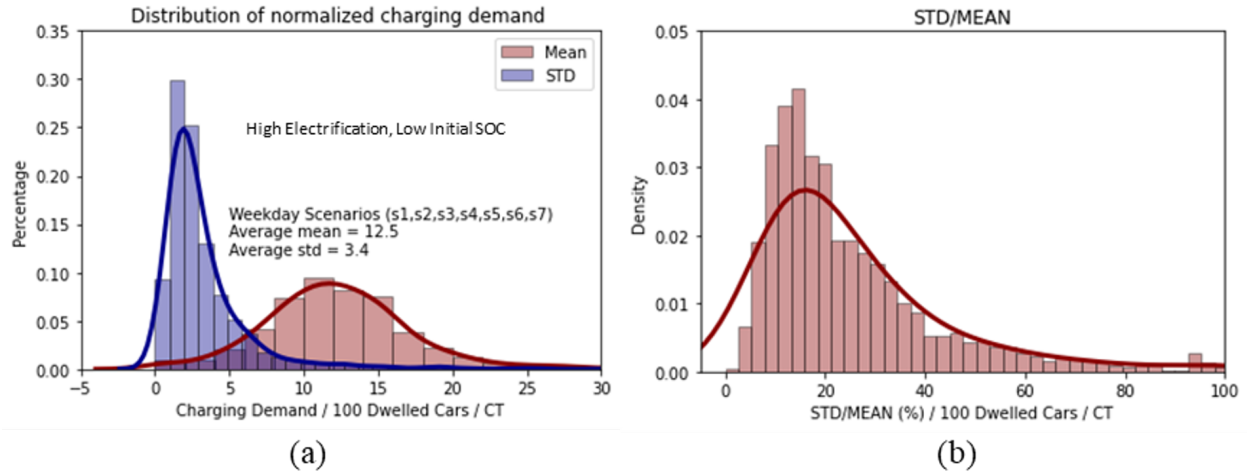


Figure 4.12: Charging demand variation across weekday scenarios: (a) Distribution of mean and standard deviation of normalized charging demand; and (b) Distribution of standard deviation divided by the mean of normalized charging demand

large. The average standard deviation of charging need mean per CT across all scenarios is around 27% of the mean of average charging need. This highlights that although the total demands generated by different scenarios are very close at the network level (see Table 4.4), the variation of charging demand across scenarios at the CT level is not negligible. This variation can be caused by the differences in the spatial and temporal distribution of incidents across scenarios, as incidents can impact link capacity, which will in turn influence the average travel speed and routing choices of EV users, and eventually affect the energy consumption rate as well as charging timing and locations.

To provide a more intuitive demonstration of how network variation impacts the charging demand of a CT, we selected one test CT near downtown LA and plotted the charging demand across 24 hours in Fig 4.13. The dashed lines in Fig 4.13(b) represent the average hourly charging demand across different network scenarios, while the semi-transparent bands indicate the lower and upper bounds of hourly charging demand across network scenarios. As shown in Fig 4.13(b), the variation in network scenarios causes significant differences in charging demand. During the peak demand hour (12 PM to 1 PM), the maximum charging



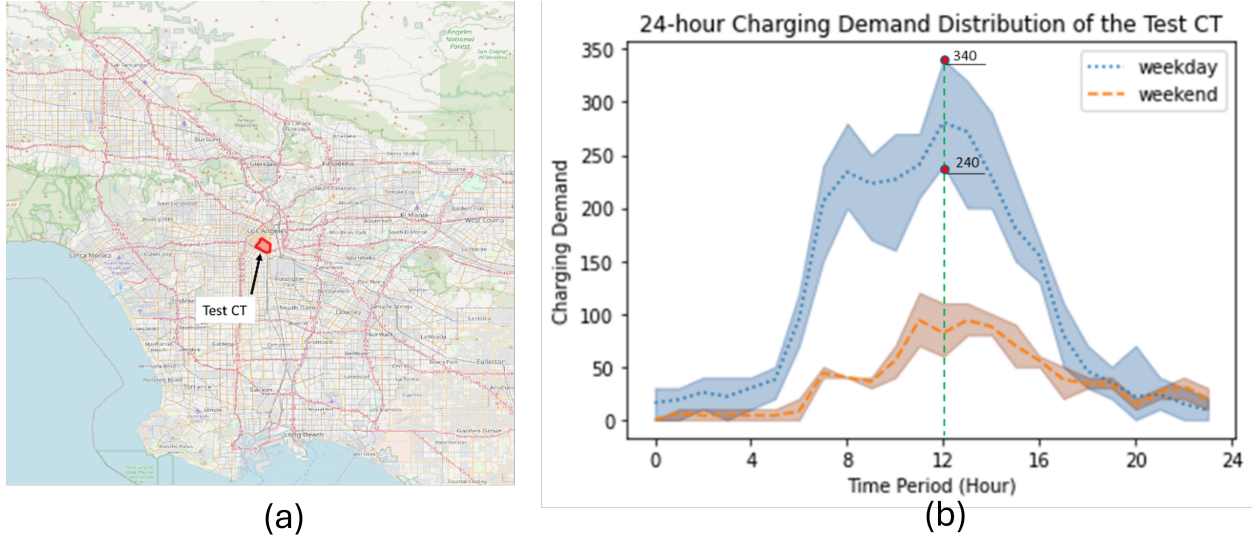


Figure 4.13: Temporal charging demand distribution of a Test CT across all network scenarios. a) The location of the test CT. b) 24-hour charging demand distribution of the test CT. The semi-transparent bands represent the upper and lower bound of the charging demand across network scenarios

demand (340) is 42% higher than the minimum charging demand (240), indicating that the variation in charging needs induced by different traffic conditions is non-negligible at the single CT level.

#### 4.4.3 Zone-level Charging Demand Variation Across Functional Areas

To explore the charging demand patterns at the zonal level, we selected sample CTs by functional area and conducted a temporal analysis for each functional area. The CTs of each functional area are defined based on the trip purpose from the SCAG ABM. For each functional area, we focus on the top 20 CTs with the largest number of trips for the selected trip function, i.e., work, school, shopping, and dining. In this section, we plot the demand comparison for each functional area under the high electrification plus low initial SOC as a sample, as the other three combinations of electrification levels and initial SOC have similar

patterns. The scenario-aggregated plots are shown in Fig 4.14, where the dashed line shows the average value and the shaded area marks the upper and lower bounds across scenarios.

As shown in Fig 4.14(a) and Fig 4.14(b), the weekday peak demand hour for working areas and school areas are close to each other, both occurring near the midday hours, around 11 AM to 1 PM. These charging demands are generated by EVs dwelling near workplaces or schools after completing commute trips. On the other hand, the peak demand hours in shopping and dining areas, shown in Fig 4.14(c) and Fig 4.14(d), occur a few hours later than in the working and school areas, from around 2 PM to 4 PM. The shift in peak demand hours in Fig 4.14(c) and Fig 4.14(d) corresponds to the fact that these non-mandatory activities (i.e., activities other than work or school) usually happen during traffic off-peak hours (e.g., 10 AM through 4 PM for LA County).

From Fig 4.14, we also find that the variation of peak demand across scenarios is more pronounced in working and school areas, where the maximum peak demand differences across scenarios can be as high as 8%. By contrast, the maximum peak demand difference in the shopping and dining areas is less than 5%. This is mainly because work and school trips often involve more long-distance travel than other types of trips, where travel speed and travel time are more likely to be affected by incidents on the routes, thereby making individual charging needs more likely to be impacted.

#### **4.4.4 Estimation of Public Chargers and Installation Cost for the Year 2035 in LA County**

The previous results discussed the scenario-based charging demand profile prediction regarding the entire network, the temporal distribution, and the spatial difference in view of diverse electrification levels, daily charging demand, and CTs. This section provides the estimation of required public chargers using the method discussed in Section 4.2.4. For a given electrification level, the required number of chargers is first estimated for each zone under each scenario and

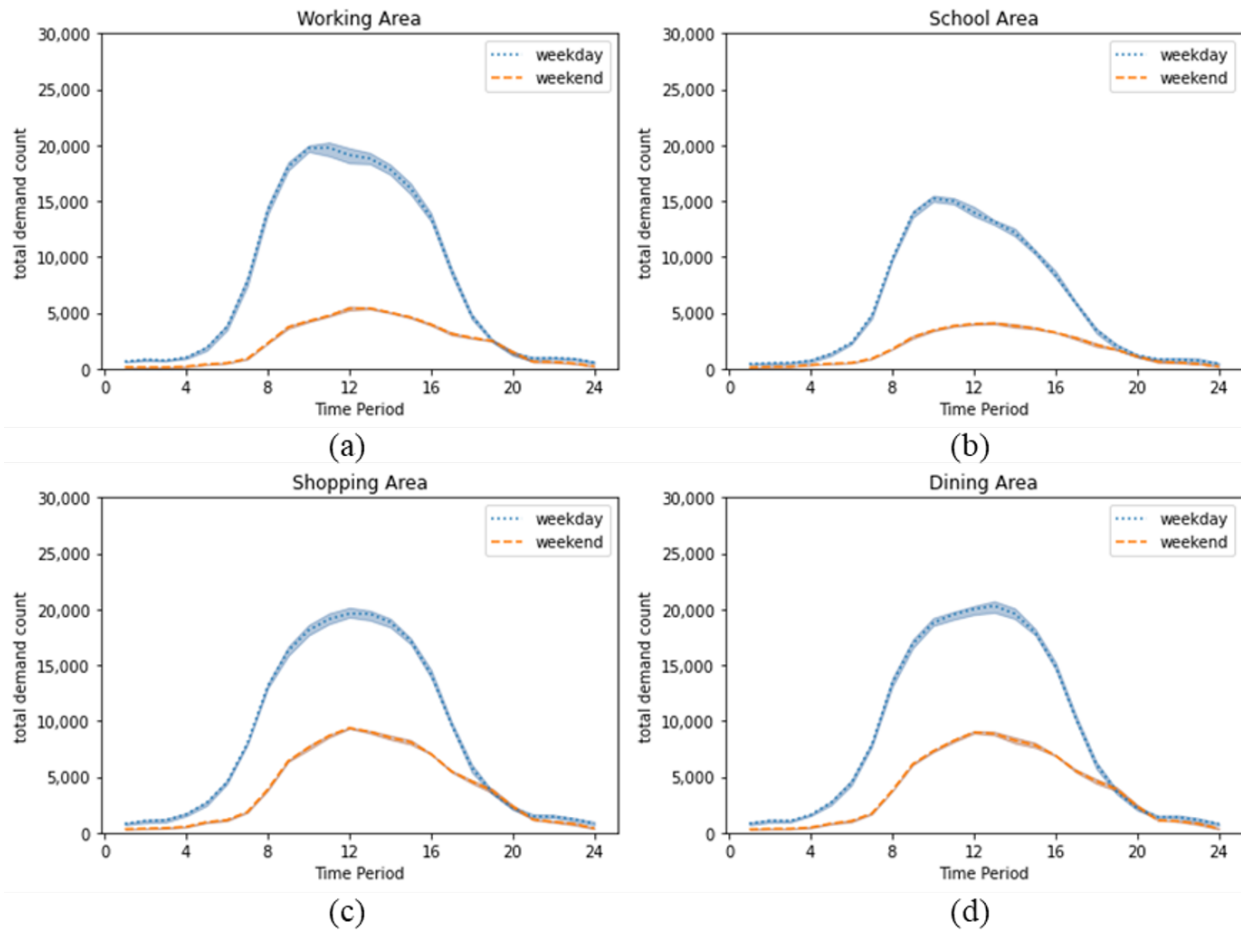


Figure 4.14: Spatial distribution of charging demand with all scenarios under high electrification level plus low initial SOC  $y$  at (a) Working area; (b) School area; (c) Shopping area; and (d) Entertaining area

then aggregated to the system level. For each zone, we use the maximum hourly required chargers across 24 hours to represent the estimated chargers for the day. Fig 4.15 provides an example of the temporal distribution of charging demand across a day for a single CT under high electrification plus low initial SOC in the network scenario S0. The CT selected for this example is the same as the test CT demonstrated in Section 4.4.2. We can observe that although the maximum hourly charging demand exceeds 300 between 12 PM to 1 PM, the maximum number of hourly required chargers is only 280, which means the required chargers to meet the charging demand of this zone under this specific scenario is 280. By repeating the aforementioned process across all zones in the system, we can determine the required chargers for the entire transportation system.

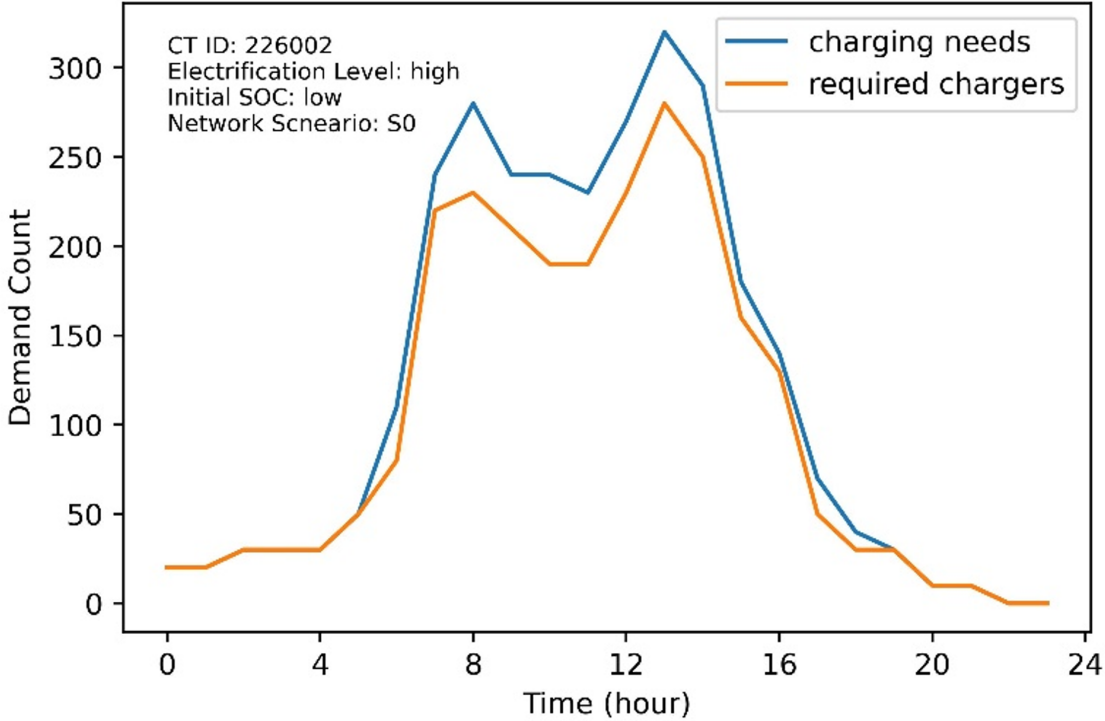


Figure 4.15: Temporal distribution of charging demand and estimated chargers for a single zone under a specific simulation scenario

With the estimation of required chargers across each zone, the next step is to estimate the

required chargers for the entire system. The estimation is conducted following the method described in Section 4.2.4 by adopting AL and SL. In this case study, due to the limited number of network scenarios, all scenarios were kept for the estimation of total required chargers. Therefore, the value of AL in this case study is set at 100%. In the meantime, we applied different levels of SL to represent the installation investment boundaries.  $SL = 100\%$  means there is no installation investment constraint, meaning the upper bound of predicted EV charging demand can all be satisfied without a waiting time between charging activities. In contrast,  $SL = 0\%$  means that the installation can only satisfy the lower bound of predicted chargers. Other SL values (25%, 50%, 75%) represent the linearly interpolated estimation of charger numbers between the upper and lower bounds for each electrification level.

Table 4.6 provides a summary of the estimated number of chargers and the corresponding installation cost under each SL for different electrification levels. This would help planners understand the cumulative investment scale by the year 2035. As shown in Table 4.6, the number of chargers reflects a significant difference depending on the electrification level. The estimated number of chargers under the moderate electrification level is only around half the number under the high electrification level, and so is the installation cost. For a given electrification level, the estimation of chargers also shows a large gap between the lower and upper bounds. The predicted L2 charger number under the high electrification level increases from 0.2 to 0.36 million; the charger number under the low electrification level ranges from 1.16 to 0.18 million.

Note that the installation cost per L2 public charger is based on the study by Smith et al. [241] and is adjustable for planners. It is expected to be cheaper in the future. Moreover, the projected number of L2 chargers is closely related to the number of plugs per L2 charger. In this case study, one L2 charger has only one plug. With the development of charging equipment, one L2 charger may have multiple plugs, which would significantly reduce the required number of public chargers.

Table 4.6: Public L2 charger predictions for LA county in 2035

<b>Electrification level</b>	<b>Unit</b>	<b>SL=0%</b>	<b>SL=25%</b>	<b>SL=50%</b>	<b>SL=75%</b>	<b>SL=100%</b>
<b>High</b>	Number	200,330	240,845	281,360	321,875	362,390
	Accumulative cost (\$M)	290.48	349.23	407.97	466.72	525.47
<b>Moderate</b>	Number	116,330	131,980	147,630	163,280	178,930
	Accumulative cost (\$M)	168.68	191.37	214.06	236.76	259.45

#### 4.4.5 Equity Assessment of Public Charging

To validate the equity benefits of improving charging opportunities through the CREATE prediction system, disparities in charger access probability across different socio-demographic groups are compared between the current year 2022 and the target year 2035. The spatial distribution for current L2 chargers is derived from current charger location data in LA County from the Alternative Fuels Data Center of the US Department of Energy (US DOE, 2022). The L2 charger spatial distribution for the year 2035 is based on the public charger prediction via CREATE, with both high and moderate electrification levels considered. For both electrification levels, we select  $SL = 50\%$  as the representative SL to indicate the public charging equity performance for the year 2035 in LA County.

Fig 4.16 shows the L2 charger access probability across different socio-demographic groups against the MFHU rate. As indicated in Fig 4.16, the charger access probability increases with the increase in MFHU rate in 2022, regardless of socio-demographic groupings. In contrast, charger access probability does not show a pronounced ascending tendency against the MFHU rate in 2035, as all MFHUs have fine access probability. We can see that even the lowest charger access probabilities in 2035 are higher than the highest access probabilities in 2022 across various socio-demographic groupings. This can be explained by CREATE considering the overall increase in public charging demand across the entire network in 2035, including disadvantaged, low-income, and minority communities.

Specifically, Fig 4.16 indicates that non-DAC, income  $> \$91k$ , and White communities have better charger access than DAC, income  $< \$44k$ , and Asian, Black, or Hispanic communities,

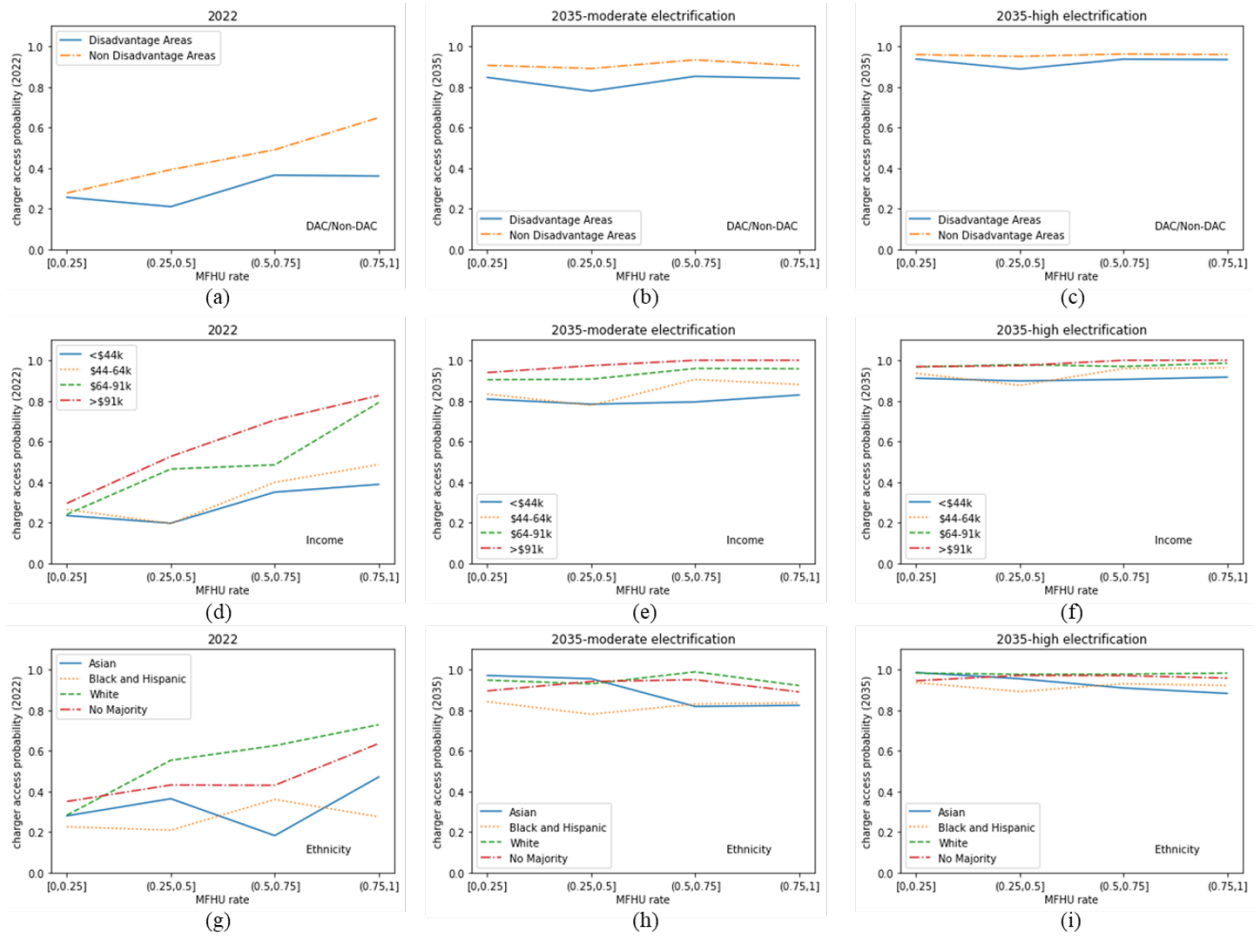


Figure 4.16: Charger access probability across different socio-demographic groups for 2022 (a, d, g), 2035 moderate electrification level (b, e, h), and 2035 high electrification level (c, f, i)

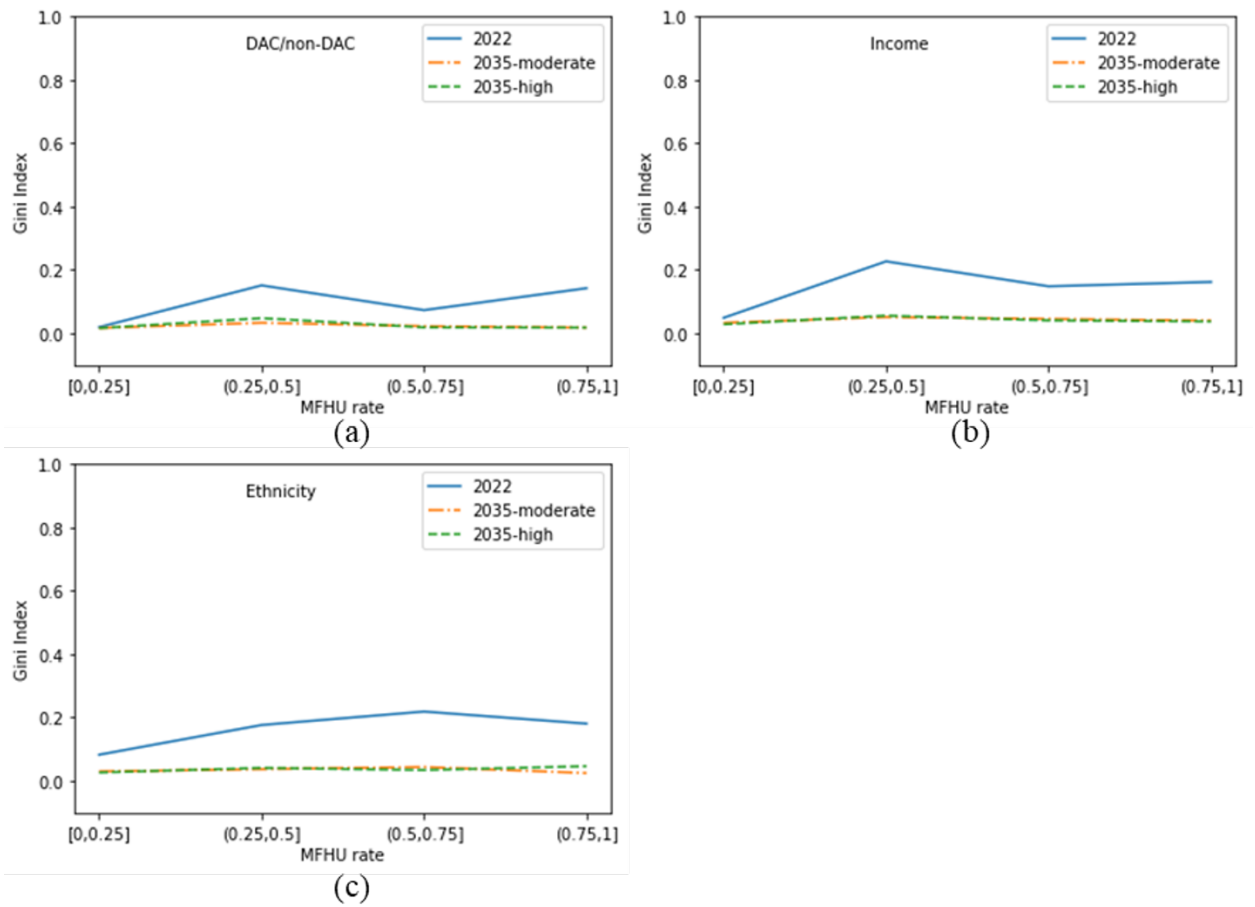


Figure 4.17: Charger access disparity of 2022 and 2035 across different socio-demographic groups: (a) DAC/non-DAC groups; (b) income groups; (c) ethnicity groups



with similar trends observed in both 2022 and 2035. The disparity in charger access probability can be attributed to disparities in EV ownership, household travel frequency, and the share of car mode in daily travel.

To assess charging equity across different socio-demographic groups, Fig 4.17 provides a quantitative comparison of the disparity in public charger access probability using the Gini index. It is evident that the charger access disparity across different socio-demographic groups decreases significantly in 2035 compared to 2022. This trend is consistent across all three socio-economic classifications under any given MFHU rates. This indicates that the disparity in charger access between communities with DAC/non-DAC, varying income levels, and ethnicities is mitigated in 2035 compared to 2022, meaning that equity in public charging accessibility is enhanced. Another finding from Fig 4.17 is that the Gini indexes for 2035 under both moderate and high electrification levels show very little difference across MFHU rates. This suggests that the two electrification levels achieve similar equity performance in terms of charger access probability.

## 4.5 Conclusion

This study developed a public charging demand profile prediction system to serve the future EV market and satisfy the future transportation system with electrification levels, which is the first study of the proposed CREATE framework. This study leveraged activity-based and agent-based transportation simulation models to generate individuals' travel trajectories and estimate each EV user's charging needs in public places. It also adopted a scenario-based approach to capture the variations in the demand and supply sides of transportation systems and to estimate the associated impacts on public charging demand.

The case study analyzed the charging demand characteristics in LA County for 2035. Simulation results indicate that:

1. The total charging demand varies significantly across different electrification and oper-

ation settings in terms of charging needs, consumed electricity, and average charging durations.

2. The zone-level charging demand reveals significant variation in zonal charging needs across network scenarios with different distributions of incidents, indicating the inevitable impact of traffic conditions on zonal charging needs.
3. The temporal charging demand pattern shows a distinguished shift in peak demand hour across functional areas, suggesting that the public charging load is correlated with people's daily activity patterns.
4. By introducing installation investment limitations, the final demand profile prediction accomplished the cost-efficiency balance between maximizing the charging service for supporting diverse electrified transportation scenarios and managing the investment under planners' financial capability.
5. Charging accessibility disparity is investigated across different socio-demographic groups to evaluate charging equity performance under the target and current years. The results indicate that charging accessibility is well enhanced in the target year in most cases, though specific communities still need additional efforts on the charging demand side to achieve a more equitable electrified transportation system.

One limitation of this study is the assumption of EV market dynamics. We use a linear projection of the EV population per zone based on 2022 EV ownership data. However, studies show that EV adoption is closely tied to charging facility accessibility [95, 217]. Additionally, EV adoption and charging infrastructure investment are interdependent [284]. Future research could employ more sophisticated models to predict the equilibrium between these two factors. Potential areas for further study include:

1. **Timing of Charging Sessions:** Improved data collection at the charging station

level would enable more accurate modeling of charging session timing and support the development of a refined charging decision model.

2. **Charger Searching Assumptions:** The current model simplified the charger searching assumption by guaranteeing chargers for every single demand. For a more realistic consideration, characteristics at charging locations such as queuing, charger selection, or the electricity load of a charging station are good research points.
3. **Integration of Charger Types:** Although fast chargers (Level 3 chargers) have strict technical requirements on the grid system that limit their wide adoption in the transportation system, it is expected to occupy 0%-10% of public chargers in the future. The next step is to combine the research domains of L2 chargers, fast chargers, and the capability of grid systems.
4. **Behavioral Modeling:** Replacing deterministic thresholds with insights from survey or charging record data could yield a more realistic representation of charging behavior.

## Chapter 5

# Deep activity model: a generative approach for human mobility pattern synthesis

Human mobility significantly impacts various aspects of society, including transportation, urban planning, and public health. The increasing availability of diverse mobility data and advancements in deep learning have revolutionized mobility modeling. Existing deep learning models, however, mainly study spatio-temporal patterns using trajectories and often fall short in capturing the underlying semantic interdependency among activities. Moreover, they are also constrained by the data source. These two factors thereby limit their realism and adaptability, respectively. Meanwhile, traditional activity-based models (ABMs) in transportation modeling rely on rigid assumptions and are costly and time-consuming to calibrate, making them difficult to adapt and scale to new regions, especially those regions with limited amount of required conventional travel data. To address these limitations, we develop a novel generative deep learning approach for human mobility modeling and synthesis, using ubiquitous and open-source data. Additionally, the model can be fine-tuned with local data, enabling adaptable and accurate representations of mobility patterns across

different regions. The model is evaluated on a nationwide dataset of the United States, where it demonstrates superior performance in generating activity chains that closely follow ground truth distributions. Further tests using state- or city-specific datasets from California, Washington, and Mexico City confirm its transferability. This innovative approach offers substantial potential to advance mobility modeling research, especially in generating human activity chains as input for downstream activity-based mobility simulation models and providing enhanced tools for urban planners and policymakers.

## 5.1 Introduction

Understanding and synthesizing human mobility patterns has become increasingly important as population growth, more complex travel behaviors, and diverse societal needs reshape modern transportation systems. Human mobility influences many facets of modern life, including traffic management, air quality, energy usage, and public health, as seen during the COVID-19 pandemic, which significantly altered travel patterns due to shifts to remote work and reduced mobility [70, 135, 260]. Other factors intertwined with human mobility include responses to natural disasters [123], traffic congestion and safety [273, 287], citizen well-being [210], air pollution [30], and energy and water consumption [171]. Consequently, human mobility modeling has garnered significant interest for its impact on these critical issues. Facing these challenges, activity-based models (ABMs) emerged in the late 1990s and early 2000s [22, 31] to capture the sequential nature of activities and predict interdependent activity choices at the individual level. ABMs are widely used by Metropolitan Planning Organizations (MPOs) in the U.S. [32]. For instance, the Southern California Association of Governments (SCAG) uses an ABM to predict activity patterns and travel demand for Southern California, a region with about 26 million people [104, 120]. Despite being state-of-the-practice models, ABMs face several limitations: data collection, model development, and calibration are costly and time-consuming; their intricate structures lead to high computational complexity; and

they rely on numerous assumptions about human activity patterns and travel behaviors, limiting their transferability to different contexts or regions.

Data-driven approaches offer promising solutions to address the limitations of ABMs. Advances in mobility data and computing power have enabled researchers to use artificial intelligence (AI) techniques for human mobility modeling. Deep learning (DL) approaches and generative algorithms replicate real-world human mobility trajectories, ensuring that generated patterns closely match observed data [18, 127, 208]. Research in generative algorithms focuses on spatial properties (e.g., travel distance distribution [88] and preferred locations [209]) and temporal properties (e.g., activity schedules [226]). Additionally, Studies show that over 90% of human activity patterns are similar across different regions, captured by a limited set of human motifs [39, 234]. In other words, a pre-trained model based on fundamental human mobility data can be fine-tuned with local data for adaptable and accurate representation in various geographic locations.

Although data-driven approaches show strong prediction capabilities, their performance relies heavily on emerging mobility data sources like GPS data, communication records, and social media from phones or cellular stations [88, 125, 139, 207]. The diverse characteristics and formats of these data sources result in unique model features that are not easily adaptable to other datasets. Additionally, some data sources are costly, difficult to access, and not widely available for diverse geographic regions, limiting the adaptability of DL models. Moreover, existing DL models mainly focus on spatial-temporal relationships using human trajectories, and often overlook the underlying semantic interdependencies of human activities, as well as the attributes of a person and their family that affect activity choices.

In this study, to address these challenges, we present a novel generative DL model, the Deep Activity model, to reveal the fundamental and generic human activity patterns that underlie human mobility of the studied region. Derived from trip diaries in household travel survey (HTS) data (Fig. 5.1 (a)), we use the concept of an "activity chain" (Fig. 5.1 (b)), which represents one-day activity sequences for individuals. HTS data forms the training

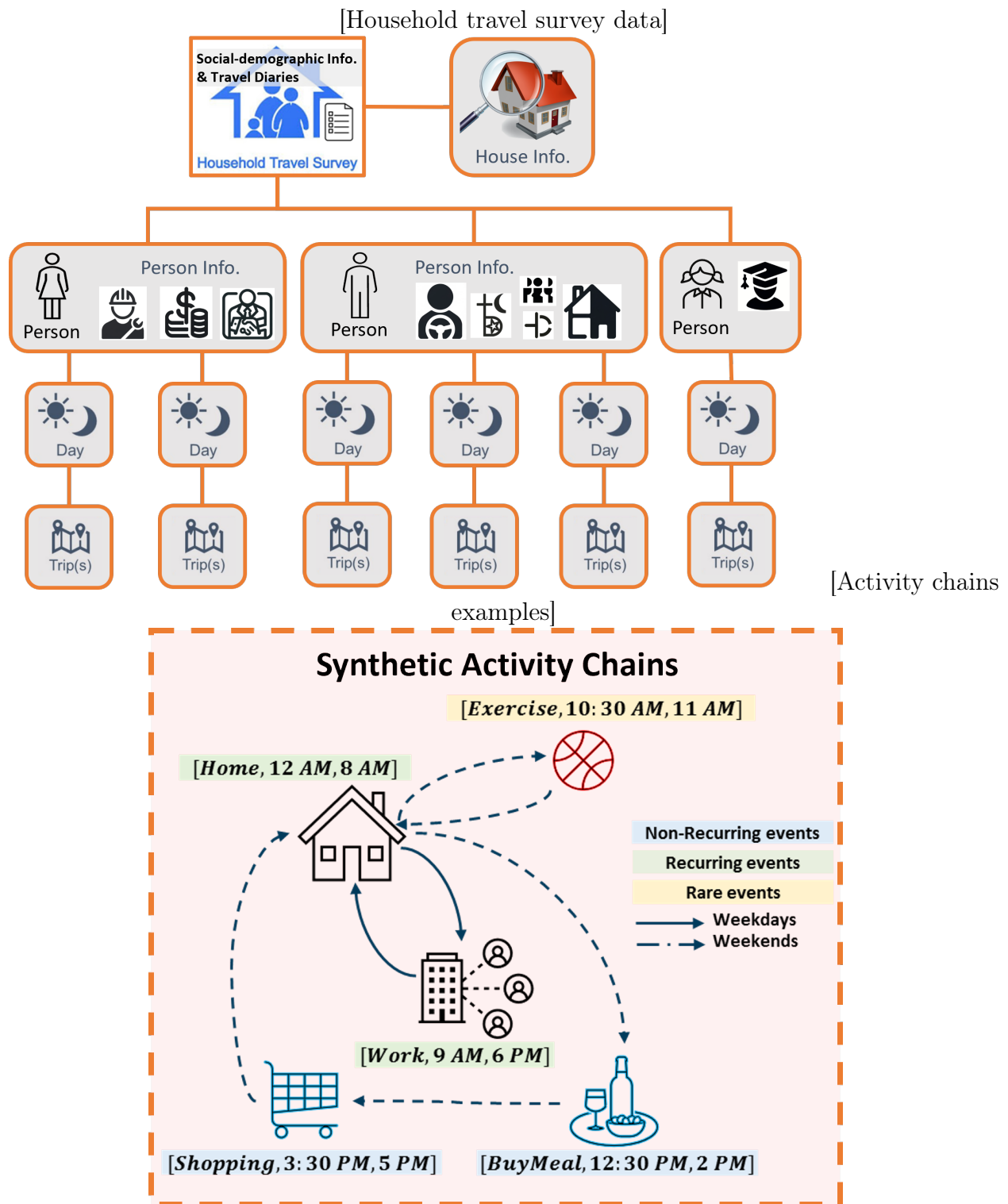


Figure 5.1: Model human mobility pattern using household travel survey (HTS) data. (a) HTS data includes information about each household member’s social-demographics, the household itself, and their daily non-commercial travel. This data covers daily non-commercial travel by all modes, along with details about the travelers, their households, and their vehicles [163]. (b) Typical weekday and weekend activity chains in HTS.

foundation for our model. Using regional population information, the Deep Activity model generate realistic and varied activity chains, effectively explaining human mobility patterns for the target region. Experiments show that the Deep Activity model generated highly realistic activity chains, evidenced by a Jensen-Shannon divergence (JSD) of just 0.001, indicating that generated chains closely match real-world data. Additionally, our model can be fine-tuned for specific regions, successfully capturing patterns in California, the Puget Sound region (in Washington), and Mexico City.

With such synthesis capability, the automatic generation of transportation system simulation models becomes much feasible. This addresses the significant challenge of the extremely high costs and labor-intensive nature of hand-crafted models that has persisted for a long time. More implications include facilitating location choice models and urban planning models by integrating the Deep Activity model. This study sets the foundation for human mobility synthesis, enabling the automatic generation of data for new regions and significantly advancing the field of transportation modeling. Our key contributions include:

- We are the first to define the human activity chain generation problem for human mobility pattern research and introduce corresponding performance metrics for evaluation. We also pioneer integrating mobility generation with activity location assignment, loading the generative travel demand within a large-scale simulation network to validate the model’s performance from a transportation system perspective.
- We propose a novel deep learning approach to generate synthetic human mobility data. Our method develops a transferable model that can be fine-tuned with local data, enabling the study of human mobility in data-limited regions by leveraging the generalizability of HTS data. By utilizing a well-designed loss function and effective input data construction, even the vanilla transformer proves highly effective in addressing this complex task.
- We explicitly model and reveal the interdependencies among activities and household



members, exploring how one household member’s decisions influence another’s.

- We explore a standard technique for multivariate, multi-objective data balancing to process the ubiquitous HTS data, pioneering its use in deep learning methods for human mobility studies. This data balancing method enhances the applicability of HTS data in modern analytical contexts and benefits the transportation modeling and planning community.

## 5.2 Related work

In recent years, human mobility has garnered significant attention due to the availability of mobility data, advancements in computing technologies, and the application of AI techniques. The literature on human mobility defines two major tasks: generation and prediction [156]. Generative models aim to create spatially and temporally realistic human trajectories. At the collective level, these models seek to replicate real-world mobility flow. On the other hand, prediction models focus on tasks such as predicting the next location for individuals [51, 289] or forecasting crowd flow distributions based on historical data [79, 261]. In this study, we specifically concentrate on generative models for human mobility patterns, particularly the activity chain generation problem.

**Transportation Models.** ABMs are the state-of-the-practice model to generate human activity patterns. An ABM is a type of modeling approach used in transportation planning and urban studies to predict and analyze individual activity patterns and travel behavior. ABMs aim to understand and simulate how people make decisions about their daily activities, such as work, shopping, education, recreation, and other social and personal activities, and how these activities influence their travel choices and travel patterns. Bowman and Ben-Akiva [31] proposed an ABM prototype to predict individual activity and travel schedule based on discrete choice models and forecasted the travel demand of the Boston metropolitan area. Goulias et al. [89] developed an ABM system (SimAGENT) to simulate the activity

and travel patterns for Southern California. The SimAGENT incorporates five components: a population synthesizer, a econometric micro-simulator for socioeconomics, a land use and transportation systems, a econometric micro-simulator for daily activity-travel, and transportation simulation [24]. Different components predict individual activity and travel choices at different levels. However, ABM’s development necessitates extensive data collection, making calibration and implementation expensive and time-consuming. Additionally, ABM’s high dependency on local data restricts its adaptability to diverse geographic regions.

**Model-Based Data Driven Methods.** Data-driven models offer an alternative approach for activity generation problems. The Exploration and Preferential Return (EPR) model, a stochastic method, simulates human mobility by balancing exploration of new locations with returns to previously visited ones [242]. In the EPR model, individuals move through a spatial environment, making decisions influenced by location popularity and distance. The model considers the balance between exploration and preferential return and evolves over time. Enhancements like TimeGeo by Jiang et al. [124] add temporal choices, such as home-based tour number, dwell rate, and burst rate, along with a hierarchical multiplicative cascade method to measure generated trips and land use. These improvements bypass HTS data limitations by offering a flexible, data-driven framework. However, model-based data-driven methods often require prior expert knowledge, and their simple implementation mechanisms may constrain realism.

**Deep Learning Approaches.** Multiple DL models were adopted to model human mobility and generate human activity and trajectories, such as fully connected networks, recurrent neural networks (RNN), attention mechanisms, convolutional neural networks (CNN), and generative models (Variational AutoEncoders (VAEs) and Generative Adversarial Networks (GANs)). See Luca et al. [156] for a detailed review. The limitations of model-based data-driven methods can be tackled by generative models (GANs or VAEs), because they can incorporate different aspects of human trajectory simultaneously (e.g., spatial and temporal features), and capture complex non-linear relationships in the data. DL model performance

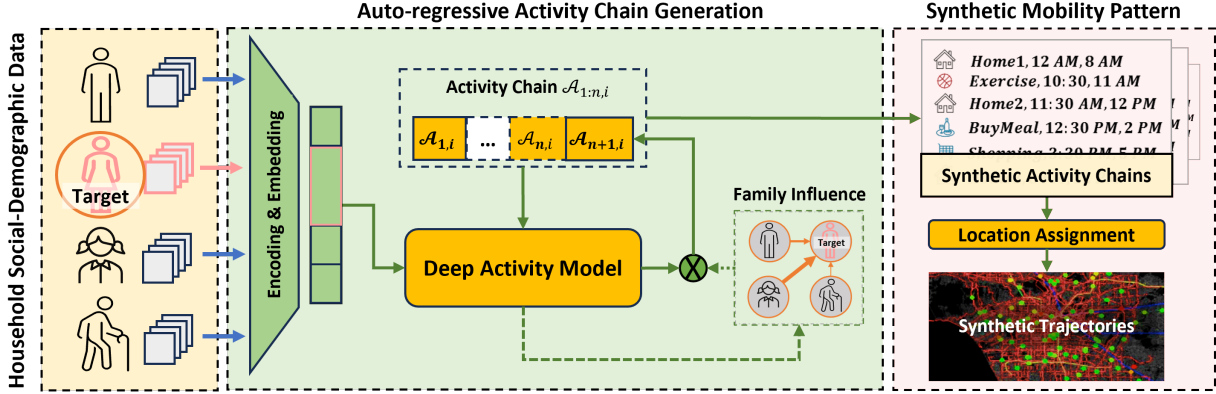


Figure 5.2: Workflow of activity chain generation

is heavily dependent on the amount and quality of data and is limited by the fact that human mobility data is usually expensive or difficult to access (e.g., requiring Non-Disclosure Agreements), as discussed in Section 5.1. Therefore, it is necessary to explore the potential of DL models to synthesize human mobility patterns using ubiquitous and open-source data.

### 5.3 Problem Formulation

One of the fundamental principles of ABM is that "the demand for travel is derived from the demand for activities" [31]. This principle highlights the nature of human mobility, where activities form the foundation of human trajectories. In this study, the concept of an "activity chain" is used to describe the structure of these trajectories, and the human mobility of a region can be represented by the activity chains of its population.

We denote  $i$  for an agent. An activity chain for the agent  $i$  is a time-ordered sequence  $A_i = \{A_{1,i}, A_{2,i}, \dots, A_{n,i}\}$ , where  $A_{n,i} = [T_{n,i}, S_{n,i}, E_{n,i}]$  represents the  $n$ -th activity conducted by agent  $i$ . Here,  $T_{n,i}$  is the activity type of  $A_{n,i}$ .  $S_{n,i}$  and  $E_{n,i}$  stand for start time and end time of  $A_{n,i}$ , respectively. Then the mobility trajectory can be expressed as  $Traj_i = \{(A_{1,i}, Z_{1,i}), \dots, (A_{n,i}, Z_{n,i})\}$ , where  $Z_{n,i}$  denotes the zone-level location where the  $A_{n,i}$  occurs. A generative model  $M$  can generate activity chains  $A_i$  for each individual  $i$ , given socio-demographic attributes of the target agent and other household members,

$D_{k,i} = \{d_{1,i}, d_{2,i}, \dots, d_{K,i}\}$ , where  $d_{K,i}$  represents the  $k$ -th socio-demographic attributes.

## 5.4 Dataset

### 5.4.1 Household Travel Survey (HTS)

To generate the activity chain for each individual in the region of interest, we rely on data collected through the HTS, following a standard format across different regions, as introduced in Fig. 5.1(a). In the US, this survey is usually conducted by federal agencies or state agencies to gather detailed information about people’s travel behaviors. The Federal Highway Administration (FHWA) administered the National Household Travel Survey (NHTS) for the United States [74]. To uncover regional activity patterns, many states also conducted statewide HTS [55, 184, 186, 191, 197, 251]. This travel-diary data source is also widely available in many countries as government agencies need such data for various purposes of public resource management. Building on this standardized HTS data, the Deep Activity model can be easily trained and transferred to other regions.

In this study, the generic model is developed using the 2017 NHTS to enhance adaptability across regions, leveraging its large dataset of over 129,600 US households, which includes demographics, activity patterns, and travel behaviors for each household member. As presented in TABLE 5.1, the activity types in NHTS are aggregated to 15 types based on the locations of activities. For instance, regular home activities and work from home are grouped as the home activity. Besides NHTS, the 2010–2012 California Household Travel Survey [198] (collected from 42,500 households), the 2017 Puget Sound Regional Travel Study [55] (collected from 3,285 households), and the 2017 Origin-Destination Survey across the Mexico City Metropolitan Area [197] (collected from 66,625 housing units) are adopted in this study for transferability exploration.

Table 5.1: Activity category code and their corresponding descriptions from 2017 NHTS

1	Home	2	Work	3	School
4	Care giving	5	Buy goods	6	Buy services
7	Buy meals	8	General errands	9	Recreational
10	Exercise	11	Visit friends	12	Health care
13	Religious	14	Something else	15	Drop off/Pick up

### 5.4.2 SCAG ABM Data

In addition to the travel survey data, we utilize synthetic mobility data [243], specifically the SCAG ABM data, for this study. The SCAG ABM dataset represents simulated human mobility patterns, providing detailed synthetic single-day activity diaries across six counties in Southern California, encompassing a population of over 19 million. The model captures 24-hour travel demand patterns at a 15-minute temporal resolution, including start and end times, types of activities, and zonal-level locations for each individual agent.

One notable advantage of the SCAG ABM data is the inclusion of zonal-level location data for each simulated activity. These zones, known as Transportation Analysis Zones (TAZs), contain demographic and spatial information about the residents and destinations within each zone, serving as both origins and destinations of trips. The use of this data significantly enhances our ability to validate the spatial-temporal performance of the mobility patterns generated by the model proposed in this study, providing a robust framework for further analysis and validation. In this study, we select Los Angeles (LA) County, which contains 5,967 TAZs, as the target area for validation.

### 5.4.3 Data Preparation

The first step in training the model involves preparing the agent information and activity chain pairs, which includes selecting relevant features for the agent information and encoding the activity chain data.

According to ABM, **socioeconomic and demographic attributes** significantly influence

individual activity patterns and travel choices. In the context of the activity generation model, the function of these attributes is similar to a prompt (as in the language model), determining the start of activity sequence generation and influencing the entire sequence generation process. These attributes include typical individual characteristics, such as gender, race, age, employment status, and job category, capturing personal and professional demographics. Education level and student status provide insights into the academic background and current academic involvement.

**Household-related** attributes such as the number of persons, relationships within the household, home ownership, and household size are also considered for their impact on daily routines and mobility. The number of vehicles owned, workers in the household, and the household employed count are indicative of transportation needs and capabilities. Household income level, along with the percentage of renter-occupied housing in the household's location, offers a socioeconomic perspective.

Additionally, **zonal attributes**, including the population density, housing units, and the classification of the residence type as rural or urban, provide a geographical context. Finally, the life cycle stage of the household is included as it reflects the evolving needs and behaviors of individuals over time. These demographic features collectively offer a comprehensive description of an individual's background. Finally, there are 13 personal attributes, 13 household shared attributes, in total 26 attributes selected to describe one individual. In conclusion, the personal and household features used for the model are presented in [Table 5.2](#).

The attribute data, originally in text or label format, is transformed into categorical data. Not everyone answers all the privacy-related questions in the survey, so for any attributes left blank or marked as 'not responded' in the NHTS, we used a dummy number for encoding. The the continuous activity start and end times are encoded with segmenting a 24-hour day into 96 intervals, each lasting 15 minutes, and numerically encoded from 1 to 96 to represent the time slots.

Regarding the **SCAG ABM dataset**, we sample subsets from the overall population of 10 million individuals for the purposes of model transfer learning and validation. Specifically, we prepare two subsets: a smaller sample of 100,000 individuals, which is used for model transfer learning, and a larger sample of approximately 1 million individuals, employed to assess the scalability of the model trained on the smaller subset to a larger population. The activity types in the SCAG ABM data are mapped to correspond to 15 categories consistent with the NHTS data, and the time variables are similarly encoded into 96 time slots, following the same procedure used for the NHTS data.

Table 5.2: Socio-demographic attributes from HTS dataset

<b>Attribute Name</b>	
Driver’s License Status	Number of Workdays
Education Level	Job Category
Gender	Age
Racial/Ethnic Identity	Weekly Transit Usage
Household Role	Household Income
Current School Grade Level	Household Size
Employment Location Type	Household Vehicles
Number of Jobs	Home Ownership Status
Employment Status	Household Students
Household Licensed Drivers	Household Life Cycle Stage
Household Employed Members	Type of Residence
Housing Status	Housing Density
Population Density	Renter-Occupied Housing Ratio

## 5.5 Methodology

The workflow for generating activity chains for an individual using our Deep Activity model is illustrated in Fig. 5.2. The process begins with household socio-demographic data, which includes information about the target person and any family members. The Deep Activity model captures the influence of household members on the target person and ensures that the generated activity chain reflects real-world interdependencies. The model then auto-

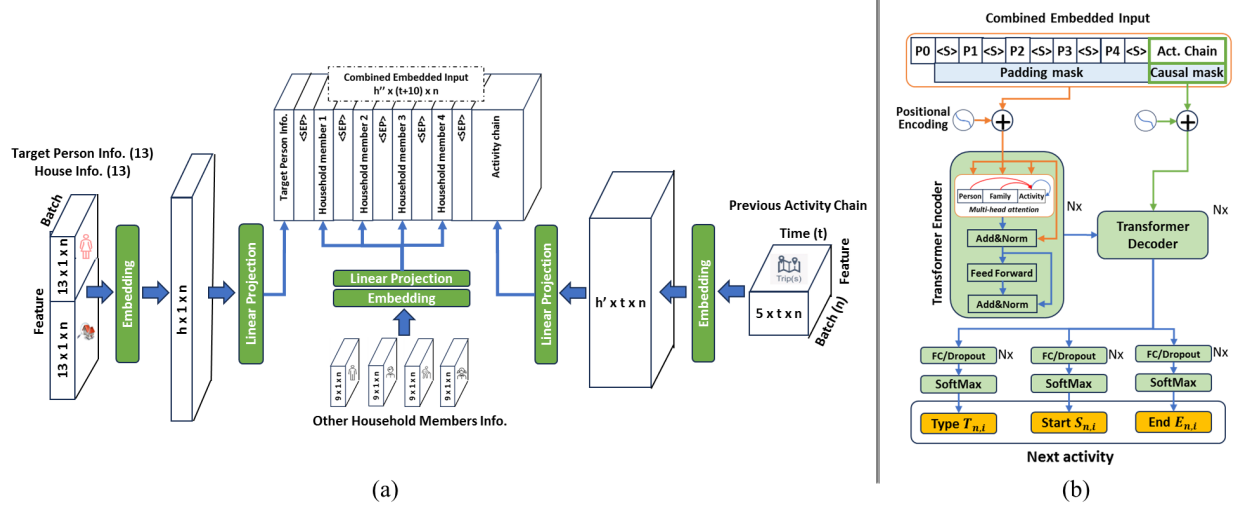


Figure 5.3: Deep Activity model architecture. (a) Input data construction. (b) Transformer-based network architecture with well-designed data injection.

regressively generates a sequence of activities, forming the target person’s activity chain  $A_i$ . Finally, location is assigned for each activity in  $A_i$ , completing the synthetic mobility trajectory generation.

### 5.5.1 Model Architectures

To generate an activity chain, comprising activity types alongside their corresponding start and end times, based on the demographic attributes of individuals. The structure of the activity chain generation problem is analogous to text generation tasks tackled by language models. Just as words in a sentence follow a logical sequence based on context, activities in a person’s daily routine are sequentially dependent on preceding activities and time constraints. Hence, a model based on Transformer [254] is developed and trained for the activity chain generation task, as shown in Fig. 5.3.

**Data structure design.** To analyze the influence of household members and previous activities on the target person’s decision, we developed an innovative data concatenation strategy. This approach combines embedded social-demographic data, data of other household members, and embedded activity data within the time domain. To standardize dimensions



across these diverse data sets, we integrated fully connected layers and employed learnable delimiters <SEP> for separation. Fig. 5.3(a) illustrates this process for a household with five members, demonstrating how the data is transformed into a comprehensive feature vector. This vector subsequently serves as input for the network’s deeper layers, enabling more nuanced analysis. Our model is designed to accommodate households of up to five members, a decision informed by statistical analysis of the NHTS dataset, which reveals that 95% of households do not exceed this size.

**Feature Embedding.** We employ embedding layers to map each categorical variable into a continuous space, utilizing the Embedding function [213], which can be optimized through backpropagation. To ensure accurate representation of categories, we created a distinct embedding layer for each categorical attribute. Formally, for a categorical feature  $c$  with  $N$  unique categories, the embedding function is defined as:  $E_c : 1, 2, \dots, N \rightarrow R^d$ , where  $d$  represents the dimension of the embedding space for that feature. The optimal value of  $d$  was determined through validation performance.

As illustrated in Fig. 5.3(a), the embedding layer processes the activity chain data, represented by a tensor with dimensions (5, t, n), where these dimensions correspond to features, time, and batch size, respectively. Additionally, another embedding layer handles the target individual’s social-demographic data, shaped as (26, 1, n), and the data pertaining to the target person and other household members. These diverse data sets are then seamlessly combined, utilizing five <SEP> delimiters to maintain clear separation between different data types.

**Network structure design.** The network structure integrates a Transformer encoder-decoder architecture. As shown in Fig. 5.3 (b), the Transformer encoder receives combined embeddings of personal and household information, and previous activities, along with padding masks to ignore irrelevant parts and causal masks to maintain the autoregressive nature of the sequence prediction. The decoder then takes the combined activity sequence and the output (memory) from encoder as additional context. By processing the entire context in

the encoder and focusing on the next activity prediction in the decoder, the model captures complex dependencies and interactions. This approach simultaneously considers personal, household, and previous activity influences on the target person’s next activity decision, improving prediction accuracy and providing a deeper understanding of the factors driving activity choices. Next, positional encoding are added to both encoder and decoder inputs to retain temporal information. Finally, the model generates predictions for the activity type  $T_{n,i}$ , start time  $S_{n,i}$ , and end time  $E_{n,i}$ , forming the next activity. The prediction process continues until either the activity marked as the end-of-the-sentence (EOS) is predicted, or the chain reaches the maximum length, at which point the prediction terminates.

### 5.5.2 Loss functions

Predicting activity type is a classification task. Because the day was segmented into 96 intervals, each lasting 15 minutes, predicting start and end times is also a classification task. **Cross-entropy loss**,  $\mathbf{L}_{\text{CE}}(\mathbf{y}, \hat{\mathbf{y}}) = -\sum_i y_i \log(\hat{y}_i)$ , is commonly used to measure the discrepancy between predicted probabilities and actual outcomes and is the first loss term to minimize activity mismatch. However, given the uncertainty of human activity, time prediction should not be overly strict but rather should allow for a certain level of deviation. Therefore, we incorporate a custom loss function that includes soft labels to allow the prediction results to deviate within a small window, enhancing the flexibility and robustness for start and end time prediction.

**Soft label loss  $\mathbf{L}_s$ .** Soft labels loss is calculated in two steps: 1) generate soft labels. 2) calculate the soft cross entropy loss. First, soft label matrix ( $S$ ) of dimension  $N \times C$  are generated by assigning higher weights to the true class and lower weights to adjacent classes.  $C$  denotes the number of activity class, and  $N$  is the batch size. For each true label step  $y_i$ ,  $S_{i,y_i} = w_m$ ,  $S_{i,y_i \pm s} = w_s$  for  $s \in [1, n_s]$ , where  $w_m$  (set to 1) and  $w_s$  (set to 0.1) are the main and side weights, respectively, and  $n_s$  is the number of allowance deviation side steps.

Then the soft label cross entropy loss for is computed as:

$$L_s = \frac{1}{N} \sum_{i=1}^N \left( - \sum_{j=1}^C S_{i,j} \log (P_{i,j} + \epsilon) \right), P_{i,j} = \frac{e^{Z_{i,j}}}{\sum_{k=1}^C e^{Z_{i,k}}} \quad (5.1)$$

where  $Z$  is the predicted result;  $P_{i,j}$  represents the probability that the  $i$ -th sample in batch  $N$  belongs to the  $j$ -th class, as predicted by the model;  $\epsilon$  is a small constant added to prevent the logarithm of zero.

Additionally, to guarantee that the generated sequence of activities adheres to a logical chronological order, two specialized time penalty losses, i.e., **temporal order loss** ( $L_o$ ) and **sequential timing loss** ( $L_{seq}$ ), are incorporated. These losses ensure that the predicted end time of an activity does not precede its start time and that the end time of a preceding activity does not exceed the start time of the subsequent activity.

$$L_o = \frac{1}{N} \sum_{i=1}^N \max (0, t_{i-1}^{\text{end}} - t_i^{\text{start}}), \quad (5.2)$$

$$L_{seq} = \frac{1}{N} \sum_{i=1}^N \max (0, t_i^{\text{start}} - t_i^{\text{end}})$$

where  $t_{i-1}^{\text{end}}$  denotes activity end time at  $i-1$  step, and  $t_i^{\text{start}}$  means activity start time at  $i$  step.

Our final loss  $L$  combines five loss terms as below:

$$L = w_1 \cdot L_{CE}(T, \hat{T}) + w_2 \cdot L_s(S, \hat{S}) \\ + w_3 \cdot L_s(E, \hat{E}) + w_4 \cdot L_o + w_5 \cdot L_{seq} \quad (5.3)$$

### 5.5.3 Data Balancing

To create a fair and representative training dataset for model development, data balancing is essential to address the imbalances in HTS data. Most people follow similar activity patterns

[39, 234], which can overshadow less common activities and lead to biased models favoring the majority class. To address this, we developed a multivariate, multi-objective data balancing technique for curating the HTS dataset. In this study, three target features need to be balanced, as shown in Table 5.3, which showcases examples of these target features, with each data sample representing a day’s activities for an individual.

---

**Algorithm 2** Data balancing for multiple target features

---

```

1: Input: (1) Training dataset, (2) Adjustment rate at each step:  $step\_size$ 
2: Output: Weights  $W$  assigned to training dataset for data resampling
3: Select  $n$  target representation features and calculate the original distributions:
    $D_{ori} = \{O_1, O_2, \dots, O_n\}$ 
4: Ideal distribution for  $n$  target features:  $D_{ideal} = \{I_1, I_2, \dots, I_n\}$ 
5: Initialize target distributions:  $D_{tar} = \{T_1, T_2, \dots, T_n\} = D_{ori} = \{O_1, O_2, \dots, O_n\}$ 
6: repeat
7:   for each feature  $i$  in  $n$  target features do
8:     // Calculate the differences between adjusted and target share for each class in
     feature  $i$ :
9:      $D_i = T_i - I_i$ 
10:     $\triangleright$  Adjust the elements based on its difference and the  $step\_size$ :
11:     $T_i = T_i - D_i \cdot step\_size$ 
12:    // Ensure the sum of percentages remains equal to 1 by normalizing the values:
13:     $F_i = 1/sum(T_i)$   $\triangleright$  define a normalization factor
14:     $T_i = T_i \cdot F_i$   $\triangleright$  update the target distribution
15:    // Calculate weights for each sample ( $W$ ) using the raking algorithm:
16:     $W = raking(D_{ori}, D_{tar}, train\_data)$ 
17:   end for
18: until raking algorithm not converging or  $|D_{tar} - D_{adj}| < threshold$ 
19: return  $W$ 

```

---

Table 5.3: Examples of target features to be balanced

id	Activity Type	Chain Length	Duration (15-min)
1	{Home, Work, Home, Exercise, Home}	5	{28, 32, 4, 7, 6}
2	{Home, School, Buy meals, Home}	4	{25, 35, 8, 24}
3	Home, Work, Home	3	{30, 60, 15}

The proposed data balancing method iteratively calculates weights for each data sample based on its significance, then performs random resampling with replacement to produce a balanced training dataset. This process can be summarized as in **Algorithm 2**, involving five key steps:

**Step 1. Feature representation:** As indicated in TABLE 5.3, activity type and duration are recorded in sequence, while the length of activity chain is a singular value, making it difficult to balance. Therefore, the most frequently occurring activity type ("mode type") and activity duration ("mode duration") are selected from each activity chain, as representations of the original target features, excluding the first and last home activities, since most of activity chains start and end at home.

**Step 2. Initial distribution  $D_{ori}$  computation:** Calculate the real class distribution for each target feature.

**Step 3. Target distribution  $D_{tar}$  specification:** Set  $D_{tar}$  as an intermediate between the actual and ideal distributions  $D_{ideal}$  (uniform distribution) to facilitate convergence.

**Step 4. Sample weights calculation:** Compute sample weights to individual samples using the raking algorithm [65], based on  $D_{ori}$  and  $D_{tar}$  from **Steps 2** and **3**.

**Step 5. Iterative refinement:** If convergence is achieved, adjust  $D_{tar}$  closer to  $D_{ideal}$  and recalculate sample weights by repeating **Steps 3** and **4**.

#### 5.5.4 Model Transfer

As aforementioned in Section 5.1, modeling human mobility patterns in regions with limited data is challenging using traditional approaches. Even with advanced deep learning methods, such as transformer models, the data-hungry nature of these models can limit their effectiveness when datasets are small [254]. By leveraging the concept of transfer learning [291], we can address this challenge effectively. The proposed Deep Activity model, initially trained on the NHTS dataset (160,000 training samples), serves as a generic pre-trained model, which

can then be fine-tuned using the limited local HTS data, adapting the generic model to the specific characteristics of smaller regions. The fine-tuning the Deep Activity model involves three primary steps:

**Step 1: Adding new layers to adapt to new features.** Augment the existing architecture with new layers to handle region-specific features and complexities, enabling better representation of diverse activity patterns.

**Step 2: Freezing part of the pre-trained model.** Initially freeze certain parts of the pre-trained NHTS model to maintain stability and leverage learned representations, preventing overfitting and ensuring effective capture of regional features.

**Step 3: Fine-tuning with regional data and unfreezing layers.** Train the modified model using regional datasets, updating weights of new layers. Gradually unfreeze selected parts of the pre-trained model, allowing comprehensive adaptation to unique regional patterns while retaining beneficial pre-trained knowledge.

The California and Puget Sound regions have input features similar to NHTS in terms of feature number, categories, and activity types. However, their datasets are relatively smaller (60,000 training samples for California and 8,000 for Puget Sound). For these regions, we apply only steps 2 and 3 of our process. In the case of Mexico City, which presents a distinct challenge due to its significantly reduced number of input features (40 compared to 60 in NHTS) and divergent activity types, we implement all three steps of our methodology.

### 5.5.5 Activity Location Assignment and Network Traffic Loading

To evaluate the Deep Activity model in a real-world transportation network and test its applicability for transportation system analysis, we propose an activity location assignment (ALA) method as an extension of the mobility pattern generation to enhance the model's functionality and ensure its practical application. This method aims to address the common limitation in travel survey data, where precise location information is often missing. The goal

is to develop a simplified location assignment method for rapid implementation in regions without high-resolution location data.

---

**Algorithm 3** Activity Location Assignment

---

- 1: **Input:** (1) Fitted distributions for all sub-regions: home-work/school distance distribution  $\mathbf{D}_{md} = \{D_{md}^1, D_{md}^2, \dots, D_{md}^n\}$ ; Non-mandatory trip distance  $\mathbf{D}_{nmd} = \{D_{nmd}^1, D_{nmd}^2, \dots, D_{nmd}^n\}$ ; location angular difference distribution  $\mathbf{D}_{ad} = \{D_{ad}^1, D_{ad}^2, \dots, D_{ad}^n\}$ ; (2) Land use types of all zones  $\mathbf{LU}$ ; (3) Zone-to-zone distance matrix  $\mathbf{M}_d$ , angle matrix  $\mathbf{M}_a$ ; (4) Home locations  $\mathbf{HL}$  and predicted activity chains  $C$  for all agents.
  - 2: **Output:** Assigned zone IDs  $\mathbf{Z} = \{Z^{md}, Z^{nmd}\}$  for all activities, including mandatory and non-mandatory activities.
  - 3: Select  $n$  target representation features and calculate the original distributions:  $D_{ori} = \{O_1, O_2, \dots, O_n\}$
  - 4: Ideal distribution for  $n$  target features:  $D_{ideal} = \{I_1, I_2, \dots, I_n\}$
  - 5: Initialize target distributions:  $D_{tar} = \{T_1, T_2, \dots, T_n\} = D_{ori} = \{O_1, O_2, \dots, O_n\}$
  - 6: **repeat**
  - 7:     **for** each agent  $i$  with mandatory activities **do**
  - 8:          $Dist_i = sample(\mathbf{D}_{md})$       $\triangleright$  Assign home-work/school distance based on agent  $i$ 's zonal distance distribution
  - 9:          $Z_i^{md} = match(HL_i, Dist_i, LU)$       $\triangleright$  Select the most matched zone based on assigned distance from work/school zones and land use type
  - 10:     **end for**
  - 11:     **for** each non-mandatory trip  $j$  **do**
  - 12:          $Dist_j = sample(\mathbf{D}_{nmd})$       $\triangleright$  Assign distance to next zone based on last activity's zonal  $D_{nmd}$
  - 13:          $Ang2Anc_j = GetAngle(\mathbf{M}_a)$       $\triangleright$  Get the angle between previous zone and next anchor zone from angle matrix
  - 14:          $AngD_j = sample(\mathbf{D}_{ad})$       $\triangleright$  Assign angle difference based on last activity's zonal  $D_{ad}$
  - 15:          $Ang_j = Ang2Anc_j + AngD_j$       $\triangleright$  Compute the direction to next location
  - 16:          $Z_j^{nmd} = match(Dist_j, Ang_j, LU)$       $\triangleright$  Select the most matched zone based on the assigned distance, angle to next zone, and land use type
  - 17:     **end for**
  - 18:     Adjust the parameters of  $\mathbf{D}_{md}$ ,  $\mathbf{D}_{nmd}$ , and  $\mathbf{D}_{ad}$  by a small margin to slightly change the shape of the distributions
  - 19: **until**  $|N_{target} - N_{assigned}| < threshold$       $\triangleright$  Stop when the error of activity numbers across sub-regions between assigned and target locations is lower than threshold
  - 20: **return**  $\mathbf{Z}$
- 

The proposed method assigns zone-level locations  $Z$  for each predicted activity by considering the distribution of distances and angular deviations between preceding and subsequent activities. For large metropolitan regions, the spatial distribution is further refined by applying sub-region-specific distance and angle distributions to capture local spatial variations.

This ensures that the assigned locations reflect the heterogeneity within different areas of the region. This is particularly important in large metropolitan areas, such as the Greater LA area, where spatial distributions may vary significantly across sub-regions, as shown in Fig. 5.4.

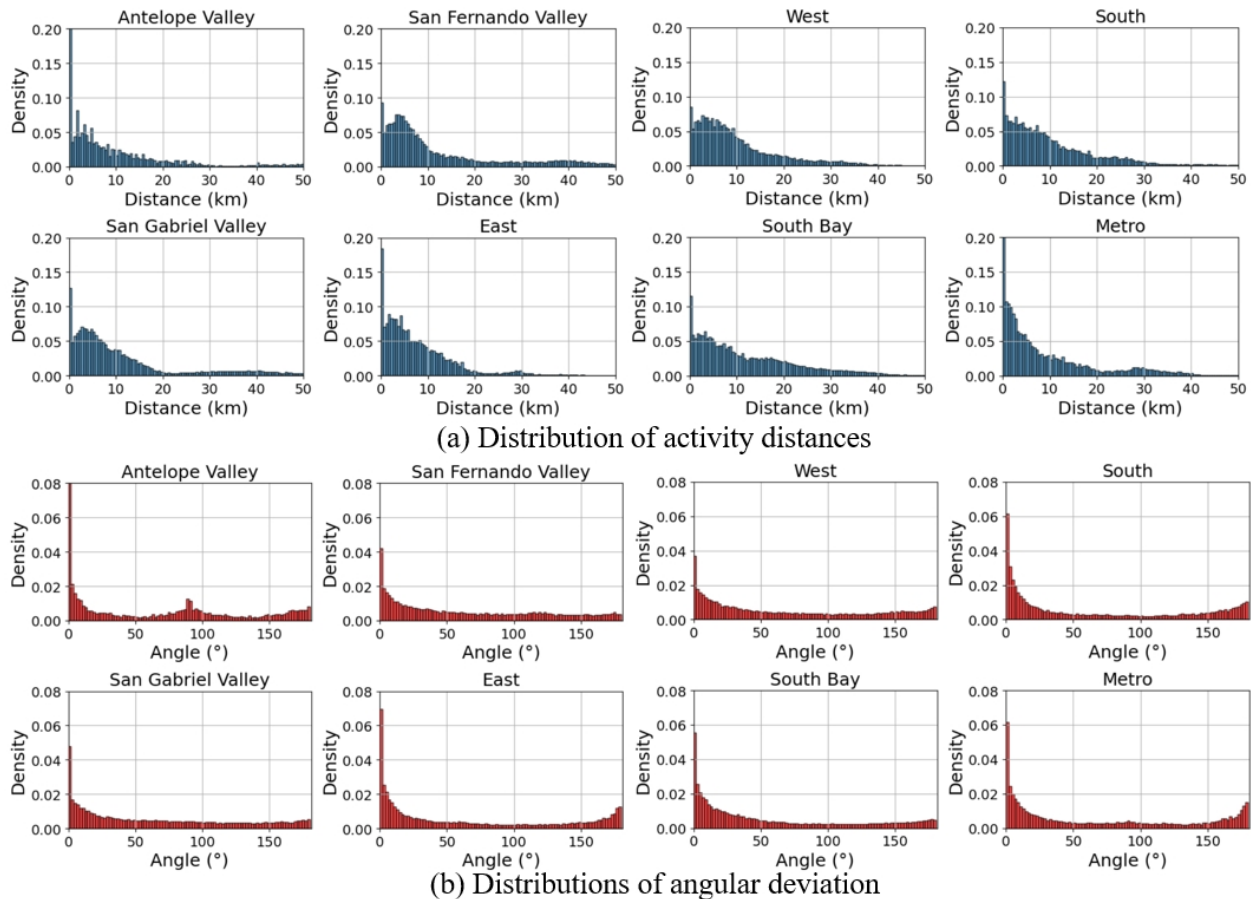


Figure 5.4: Distributions of activity distances and angular deviations across sub-regions in LA

The proposed ALA process consists of the following key steps, as shown in **Algorithm 3**:

**Step 1. Assigning TAZs for Mandatory Activities.** The primary assumption is that each individual’s home location is predetermined in the dataset. The first group of activities to be assigned locations are mandatory activities (e.g., work or school). For each individual, a home-to-work or home-to-school distance is assigned based on their demographic



characteristics. Within each sub-region, these assigned distances follow the target distribution of commute distances between home and mandatory activity locations for that specific sub-region, denoted as  $D_{md}$ . A TAZ with the appropriate land use type (work or school) and closely matching the assigned commute distance is then allocated to the individual.

**Step 2. Assigning TAZs for Non-Mandatory Activities.** After assigning TAZs for mandatory activities, these locations serve as anchor points within the individual’s daily activity chain. The subsequent step involves assigning locations for non-mandatory activities (e.g., shopping, leisure, exercise) that occur between these anchor points. Non-mandatory activity locations are assigned based on two key parameters: (1) the distance to the next non-mandatory activity, and (2) the angular deviation between the direct path to the next non-mandatory location and the direct path to the next anchor location. The assigned distances and angular deviations follow the target distributions of distance ( $D_{nmd}$ ) and angular deviations ( $D_{ad}$ ) for each sub-region, ensuring that the spatial distribution aligns with regional patterns.

**Step 3. Refinement of Location Assignment.** The objective of the location assignment process is to ensure that the spatial distribution of generated activity locations closely resembles the true activity distribution across sub-regions. This spatial similarity is assessed by comparing the occurrence frequencies of activities across sub-regions in the location assignment output with those in the ground truth data. To minimize bias in the generated distribution, the reference distributions used in Steps 1 and 2 are iteratively adjusted until the assigned activity number of sub-regions match the ground truth. This refinement process is initially applied to a small sample of the population to fine-tune the reference distributions for each sub-region. The refined distributions ( $D_{md}$ ,  $D_{nmd}$ , and  $D_{ad}$ ) are then applied to the larger population for large-scale transportation system analysis.

It is important to note that the objective of this location assignment method is not to predict the exact location of each activity, but rather to ensure a realistic spatial distribution of activity locations across the study area. This method provides a suitable input for subsequent

transportation system analyses, such as regional traffic volume estimation and congestion assessment.

The mobility patterns and the corresponding activity location assignments together form a comprehensive regional travel demand input. This generated travel demand is integrated into an existing transportation simulation framework, LASim [103]. LASim is a large-scale, agent-based multimodal transportation simulation designed for the Greater LA area, as shown in Fig. 5.11 (c). The framework builds on the Multi-Agent Transport Simulation (MATSim) to tackle the challenges induced by urbanization and changing mobility patterns. By loading the synthesized travel demand into the LA roadway network, we can generate the synthetic traffic flow and further evaluate the transportation system performance based on the generated human mobility patterns.

## 5.6 Experiments and Results

### 5.6.1 Training and Evaluation Methods

All experiments are conducted on a NVIDIA RTX A5000 GPU. We employ the Adam optimizer with an initial learning rate of 0.005. The scheduler multiplicative decays the learning rate by a factor of 0.95 after each epoch. The model is trained on a dataset with 160,831 activity chains for training and 18,106 for validation over 150 epochs with a batch size of 512. To prevent overfitting, we use regularization methods like dropout and early stopping. Finally, a test set of 18,106 activity chains is used to evaluate performance.

Given the inherent uncertainty in human behavior, evaluating the accuracy of a specific agent can be challenging and may not always be appropriate. Therefore, the performance of the Deep Activity model is assessed at the system level by comparing the similarity between the distributions of generated and real-world activity patterns, and the location assignment is evaluated at the traffic network level.

In this study, the Jensen-Shannon Divergence (JSD) is used as the similarity metric [156], as shown in Equation 5.4. The goal is to minimize the difference between the distributions of generated and real activity patterns from activity chains. The metrics include: 1) **activity frequencies**, 2) **start times**, 3) **end times**, 4) **number of daily activities** (activity chain length), and 5) **duration of each activity**.

$$\begin{aligned}
 JSD(P\|Q) = & \frac{1}{2} \sum_{x \in X} \left[ P(x) \log \left( \frac{P(x)}{M(x)} \right) \right] \\
 & + \frac{1}{2} \sum_{x \in X} \left[ Q(x) \log \left( \frac{Q(x)}{M(x)} \right) \right]
 \end{aligned} \tag{5.4}$$

where  $M = (P + Q)/2$ . Here,  $P$  is the distribution of activity patterns from the generated activity chains, and  $Q$  is the distribution from the ground truth activity chains.  $X$  represents the full range of probabilities for a specific activity pattern statistic. A JSD value closer to zero indicates greater similarity between the distributions, showing the model’s effectiveness in approximating the true distribution.

In addition to quantifying the model’s performance at the distribution level using JSD values, it is also critical to analyze it at the chain level. By aggregating activity chains into graphs, where activity types are nodes and transitions between activities are edges, **completeness** of the activity chains, as the sixth metric, can be examined to ensure no activity types or transitions are missing. These graphs can be converted into transition matrices, which reveal the transition probabilities between any two pairs of activities. As the seventh metric, the **similarity of activity transition probability** is quantified by **Frobenius norm**, i.e.,  $|A - B|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij} - b_{ij}|^2}$ , where matrix  $A$  and  $B$  are the transition matrix of generated activity chains and the ground truth, respectively.

For the validation of the ALA and network traffic loading, given that the objective of the location assignment is not to precisely predict the next location in each activity chain, we do

not employ metrics typically used in location prediction tasks, such as the accuracy of top-k recommended locations [151]. Instead, we compare the performance of the proposed method with SCAG ABM using transportation system-level metrics, including the cosine similarity of Origin-Destination Flow matrices [92], the number of activities across sub-regions [178], hourly vehicle-miles-traveled (VMT) [120], traffic volume [120], and traffic speed [120] over a 24-hour period. The error is quantified using Mean Absolute Percentage Errors (MAPE), a widely adopted measure in transportation system analysis [144]. We also conduct performance comparisons at multiple scales, including both the whole network level and a selected corridor level.

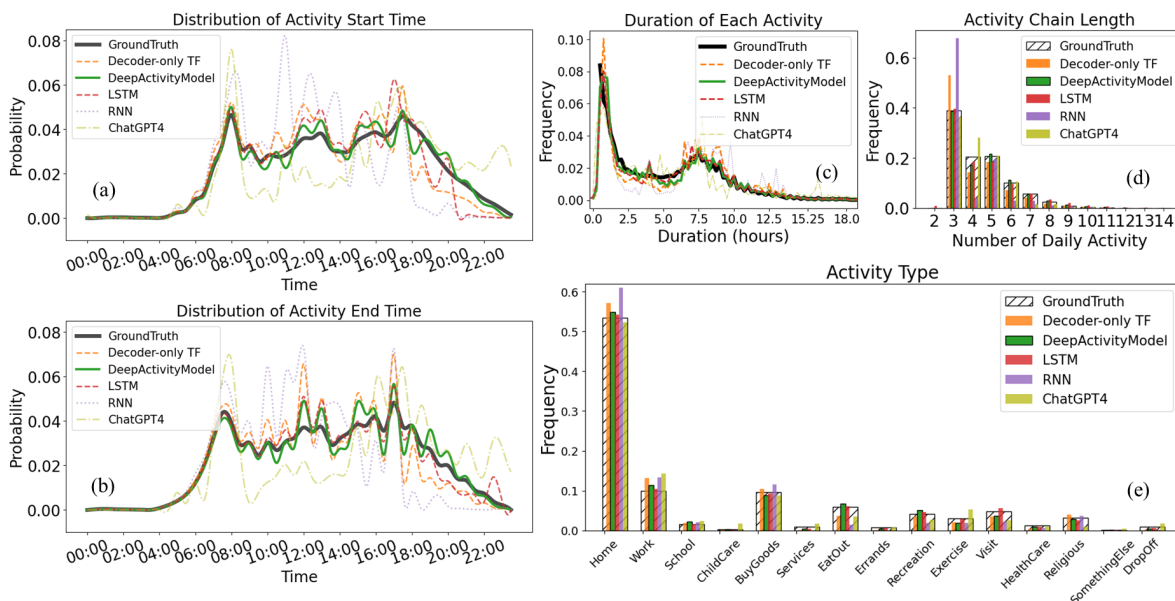


Figure 5.5: Detailed analysis and comparison of activity generation on (a)(b) temporal dynamics, (c) activity chain length, (d) activity duration, and (e) activity type distribution.

## 5.6.2 Baseline Models

**Decoder-only transformer.** In language modeling and sequence generation tasks, the "decoder-only transformer" is commonly adopted due to its effectiveness [227], making it a natural baseline for activity generation, allowing for a clear comparison when evaluating more complex models.

Recurrent Neural Network (RNN) and their variants are widely used for predicting travel behavior, e.g., next location prediction [13]. In this study, they are used as a comparison against the transformer-based models.

**Vanilla RNN** has a simple architecture where the output from the previous step is fed back into the network to influence the output of the current step [13].

**Gated Recurrent Unit (GRU)** introduces gating mechanisms to control the flow of information, maintain long-term dependencies, and address the vanishing gradient problem [146].

**Long Short-Term Memory (LSTM)** features a more complex architecture, consisting of three gates: the input gate, the forget gate, and the output gate [136], compared to the two gates used by GRU.

**Large Language Models (LLMs)** have demonstrated exceptional capabilities in understanding context and generating complex sequences without the need for extensive training periods, making them a suitable baseline for human mobility modeling. In our previous study [155], we utilized pre-trained models such as ChatGPT-4 and the open-source Llama2-70b to generate daily activity chains based solely on socio-demographic information, without the need for long-term training on domain-specific data.

### 5.6.3 Evaluation on Activity Generation

#### Distribution Similarity

To evaluate the performance of the proposed Deep Activity model, a comparative analysis is conducted involving the baseline models. The results of seven metrics are presented in Table 5.5, where the proposed Deep Activity model outperforms the others by achieving the lowest JSD values for activity chain length, duration, start time, and end time. These results indicate a high degree of similarity between the generated and ground truth activity patterns, underscoring the model’s accuracy in capturing dynamic human activities. Additionally,

it excels in edge completeness, with a percentage of 92.2%, significantly surpassing other models and illustrating its robustness in capturing the full spectrum of activity transitions. The LSTM model stands out in two metrics, showing superior performance compared to the Decoder-only Transformer. It achieves the lowest JSD value for activity type and the lowest Frobenius norm, indicating minimal discrepancy in transition probabilities between the generated and ground truth activity chains.

The LLMs demonstrate mixed results. While LLMs like GPT-4 show promise in activity type and chain length prediction, their higher JSD values for temporal aspects and lower edge completeness scores reveal significant limitations in capturing the complex dynamics of daily activities. These constraints, particularly evident in LLaMA2’s underperformance across all metrics, indicate that current LLMs are not yet suitable for generating accurate activity chains without substantial adaptations to better model the nuanced patterns of human routines.

On the other hand, traditional models like GRU and RNN, with notably higher JSD values, demonstrate moderate performance but fall behind more advanced approaches. This indicates that they struggle to capture the complex temporal dependencies and transitions that characterize human activity patterns.

Table 5.4: Activity chain generation evaluation

Model	Len.	Dur.	Start	End	Type	EC	F-Norm
GRU	0.015	0.032	0.085	0.217	0.013	51.6%	0.934
RNN	0.064	0.029	0.067	0.121	0.024	50.9%	1.291
LSTM	0.006	0.003	0.019	0.004	<b>0.003</b>	91.4%	<b>0.377</b>
D-TF	0.011	0.011	0.014	0.012	0.013	67.1%	0.784
GPT-4	0.011	0.018	0.064	0.074	0.009	42.4%	1.111
LLaMA2	0.048	0.024	0.159	0.156	0.045	19.9%	1.404
Proposed	<b>0.002</b>	<b>0.002</b>	<b>0.003</b>	<b>0.003</b>	0.005	<b>92.2%</b>	0.643

Table 5.5: \*

\*D-TF: decoder-only transformer; Len.: activity chain length; Dur.: duration of each activity; EC: edge completeness in percentage; F-Norm: Frobenius norm. All models reach 100% node (activity occurrence) completeness. Numbers except EC and F-Norm are JSD values.

In addition to the quantitative analysis, Fig. 5.5 provides a deeper insight into how

each model performs in predicting specific details of human activities. Poorly performing models were excluded from the figure to maintain focus on the most relevant comparisons. For instance, in terms of start times (in Fig. 5.5(a)), both the LSTM and Decoder-only Transformer underestimate evening activities, whereas for end times (in Fig. 5.5(b)), the LSTM overestimates evening activities, and the Decoder-only Transformer overestimates midday activities. Meanwhile, ChatGPT4 demonstrates high volatility in predicting both start and end times, suggesting challenges in capturing consistent temporal patterns. Regarding the activity chain length (in Fig. 5.5(c)), the Decoder-only Transformer tends to generate three activities per day for individuals, while all models, except ChatGPT4, tend to underestimate the four-activity chains. This suggests a tendency in most models to simplify daily activity sequences, potentially missing the complexity of real-world behavior. In terms of activity types (in Fig. 5.5(e)), Transformer-based models, LSTM, and ChatGPT4 perform well, especially for the common activities.

### Loss Term Ablation Study

To assess the impact of individual loss terms in our Deep Activity model, we performed an ablation study, summarized in Table 5.6. The findings illustrate how each loss term contributes to model performance across the seven metrics.

Removing the **soft label loss** ( $L_s$ ) has the most significant impact across all metrics. This underscores the critical role of  $L_s$  in encouraging the model to explore different combinations and learn overall temporal patterns, rather than overfitting to specific time stamps. The flexibility provided by  $L_s$  appears crucial for capturing the inherent variability in human activity schedules. The absence of the **temporal order loss** ( $L_o$ ) results in noticeable performance drops across all metrics, including duration and activity chain length. While the impact on start and end time predictions is less severe than removing  $L_s$ , the decline in length accuracy suggests that  $L_o$  plays a role in maintaining not just the chronological order, but also the overall structure of daily activity chains. When the **sequential timing**

**loss** ( $L_{seq}$ ) is removed, we observe relatively minor decreases in most metrics, with duration accuracy remaining largely unchanged. However, the higher Frobenius norm indicates that  $L_{seq}$  is particularly important for maintaining accurate activity transition probabilities, even if its impact on individual activity timings is less pronounced.

Table 5.6: Influence of each loss term

<b>Loss</b>	<b>Len.</b>	<b>Dur.</b>	<b>Start</b>	<b>End</b>	<b>Type</b>	<b>EC</b>	<b>F-Norm</b>
w/o $L_s$	.025	.012	.039	.024	.018	65.6%	.907
w/o $L_o$	.024	.011	.017	.016	.014	68.5%	.711
w/o $L_{seq}$	.013	.006	.015	.014	.013	72.6%	.739
All	.002	.002	.003	.003	.005	92.2%	.643

## Contextual Variation

Distinct activity patterns between weekdays and weekends are effectively captured by the proposed model, as presented in Fig. 5.6. The start time distribution (a) reveals a later peak for weekend activities compared to weekdays, with a notable weekend shift towards midday starts. In Fig. 5.6(b), there are more activities around 7.5 hours during week, implying the working hours, which can also be reflected in Fig. 5.6(c). Clear variations are displayed in activity types, with work-related activities dominating weekdays while leisure activities like "EatOut" and "Visit" increase on weekends. Notably, the proposed model closely mirrors these temporal and categorical differences, demonstrating its ability to distinguish and reproduce weekday-weekend variations in human activity patterns.

Age is another crucial factor influencing activity patterns synthesis process, alongside day-of-week. Fig. 5.7 illustrates how the proposed model captures age-related differences across young (0-18), middle-aged (19-65), and elderly (65+) groups. In Fig. 5.7 (a), distinct end-time distributions are observed. Young individuals show peaks aligned with school schedules, middle-aged adults exhibit a more varied pattern reflecting diverse work commitments, while the elderly display a gradual curve peaking around midday. The activity type distribution in



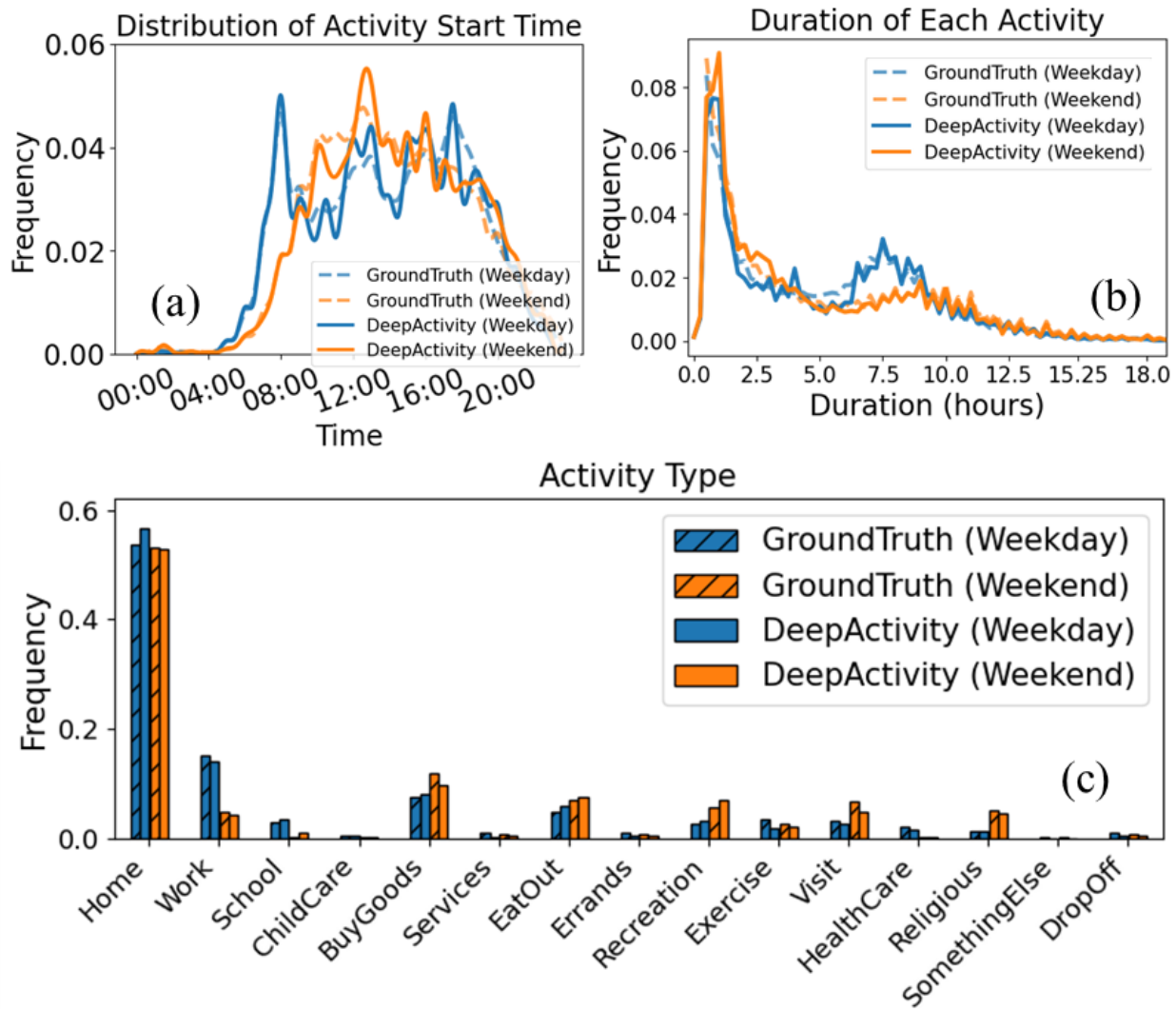


Figure 5.6: Activity patterns in weekdays and weekends

Fig. 5.7 (b) further highlights these differences, with high school attendance and recreation for the young, significant work-related activities for middle-aged, and increased shopping and eating time for the elderly. Notably, the Deep Activity model accurately reproduces these age-specific patterns in both timing and activity types, demonstrating its ability to synthesize realistic activity chains that reflect the distinct lifestyles associated with different age groups.

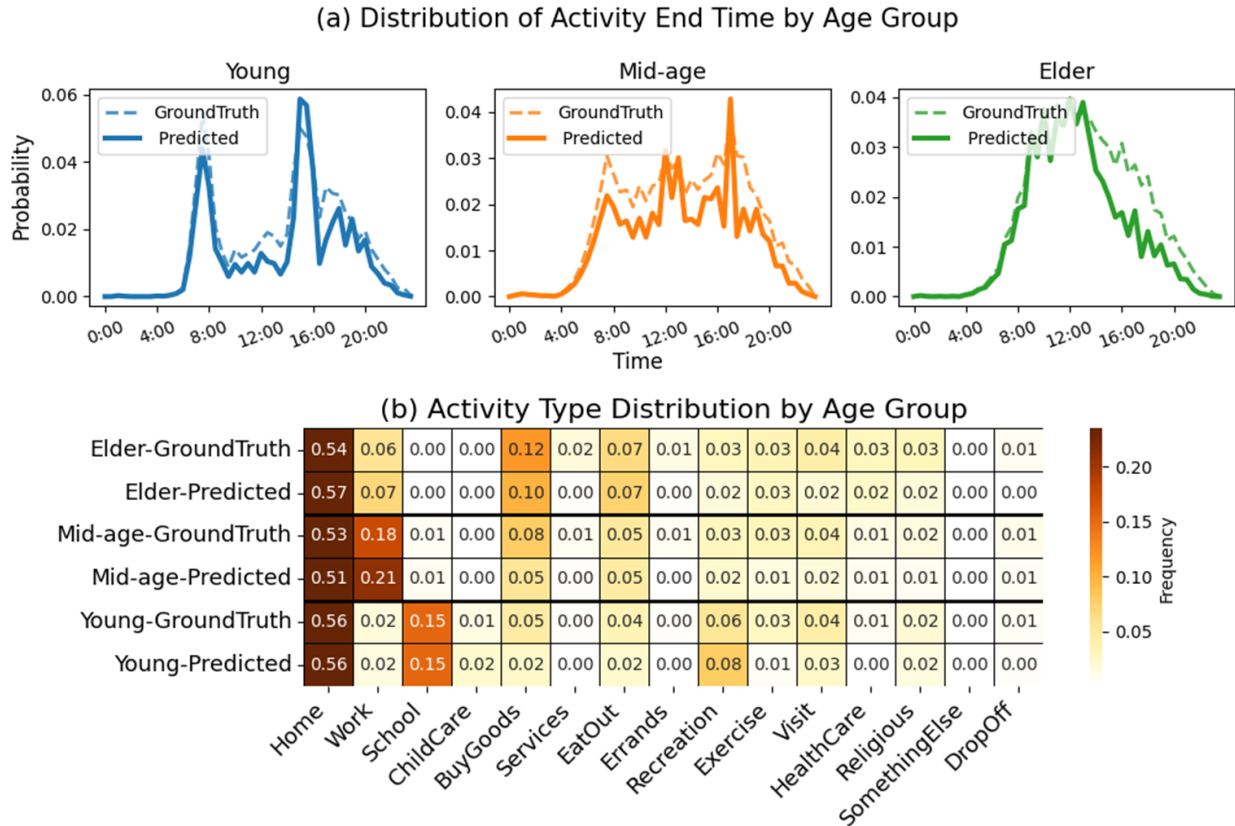
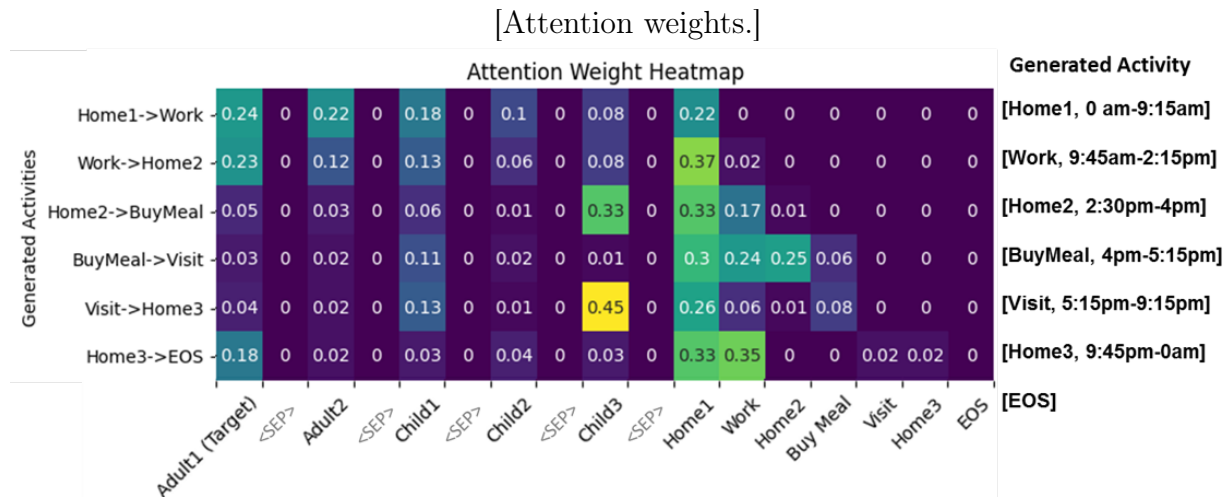


Figure 5.7: Activity patterns across age groups

## Interdependency among Household members and Activity in Activity Generation

The utilization of attention mechanisms in transformer models provides valuable insights into the decision-making processes within families, as illustrated by the detailed attention heatmap and graphical representation in Fig. 5.8. With its unique input data design, we are able to visit the attention from the first layer of encoder of the Deep Activity model, which reveals the interdependencies among the person’s household and their activities.

Each row in the heat map (Fig. 5.8(a)) corresponds to specific activities of the target person—a male worker and father in a five-member family. The columns demonstrate how interactions with other family members and previous activities influence subsequent activities. The step-by-step activity generation is detailed in Fig. 5.8(b), where each step is labeled in a unique color, and the varied thickness of the lines indicates the relative influence of each



[Activity chain generation process.]

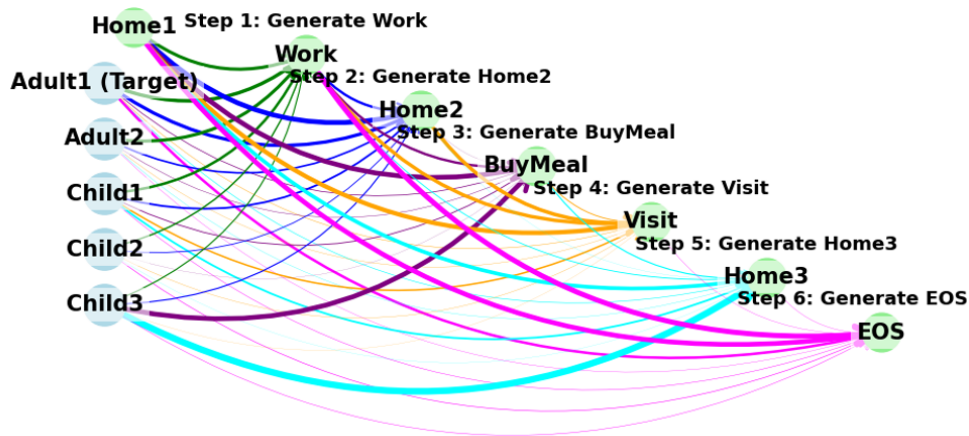


Figure 5.8: Attention weights reveals the interdependency among household members and activities.

interaction. For example, Child 3 exerts a significant influence on the target person's activities, such as "BuyMeal" and "Visit," highlighting the interdependencies of family members in coordinating daily schedules.

## Model Transferability

To demonstrate the transferability of the proposed model, we applied fine-tuning techniques to three diverse regions: California, Puget Sound, and Mexico City. These areas represent a wide range of different geographies and sizes, each exhibiting unique mobility patterns, as shown in Fig. 5.9. The datasets highlight distinct variations in activity timing, duration,

type, and daily frequency.

For activity end times, Puget Sound and California peak around 17:00-18:00, while Mexico City shows three peaks, notably at 08:00 and 14:00-15:00. Puget Sound has more short activities (less than 0.5 hours) compared to the other regions. Chain lengths vary, with Mexico City having more 3-activity chains than other regions. "Home" is the most frequent activity type across all regions, followed by "Work." Mexico City has a higher proportion of "Home" and fewer "Work" activities than the U.S. regions. Activity types like "EatOut," "ChildCare," and "Visit" present in the U.S. datasets are missing from Mexico City HTS, where "Exercise" is grouped under "Recreation." These differences highlight challenges in standardizing activity classifications, emphasizing the need for region-specific fine-tuning of the NHTS pre-trained model to capture unique mobility behaviors.

Our fine-tuning approach has proven highly effective, as evidenced in Table 5.7. The fine-tuned models successfully adapted to different regions, maintaining robustness and achieving accuracy levels comparable to those derived from the full NHTS dataset.

Table 5.7: Transferred Deep Activity model to other regions

<b>Region</b>	<b>Len.</b>	<b>Start</b>	<b>End</b>	<b>Dur.</b>	<b>Type</b>	<b>EC</b>	<b>F-Norm</b>
California	.004	.013	.033	.002	.007	83.6%	.460
Puget Sound	.030	.009	.051	.010	.012	79.5%	.652
Mexico City	.010	.056	.010	.006	.009	63%	.339

To further illustrate the performance of our proposed method, we visualized detailed fine-tuning results for each region, as illustrated in Fig. 5.9. Importantly, although the transfer learning only use very few samples, the model’s predictions closely match the ground truth across all metrics and regions, with only minor discrepancies. This demonstrates the model’s effectiveness in capturing and reproducing diverse regional mobility patterns, validating the success of the transfer learning approach in adapting to different urban environments.

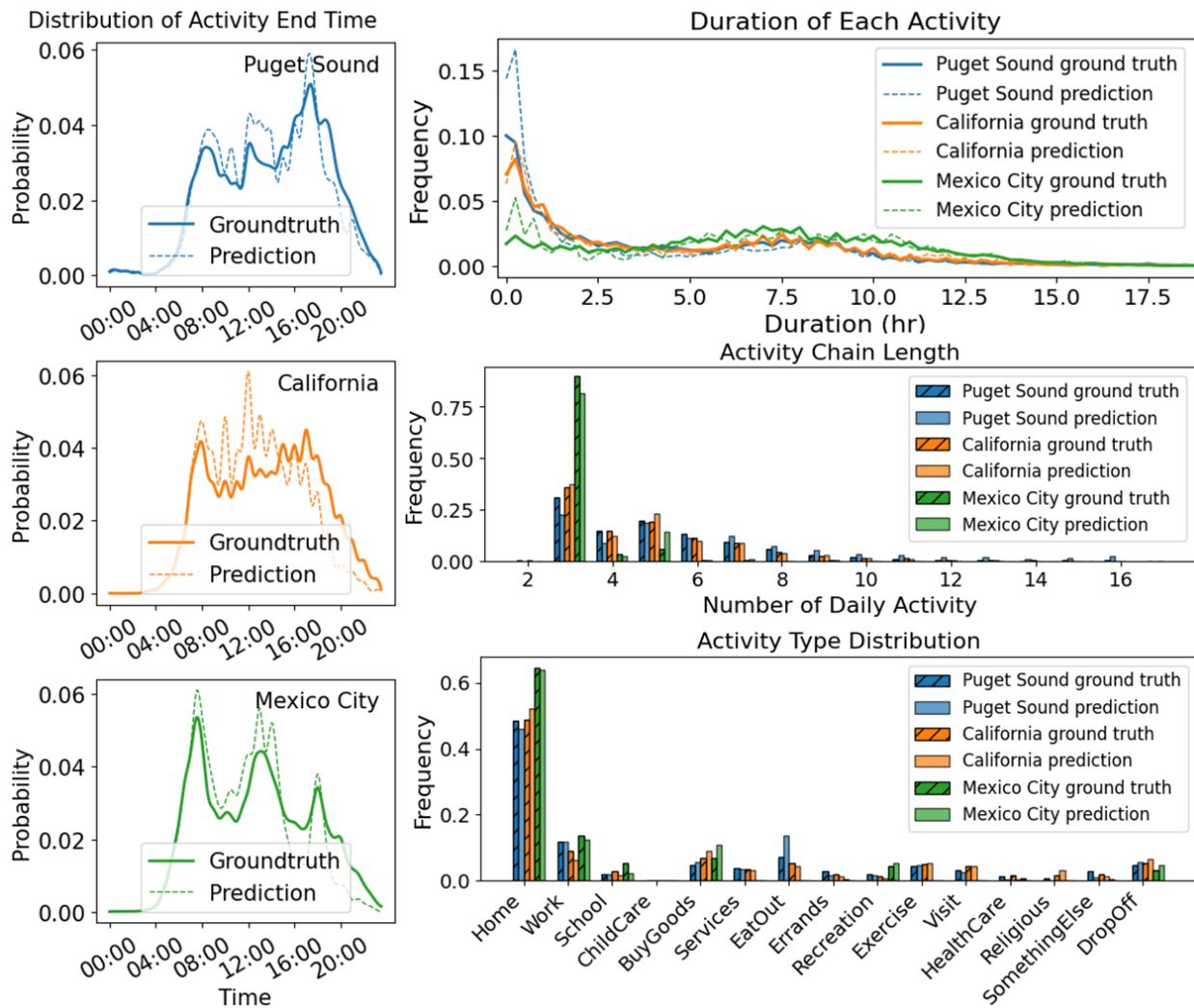


Figure 5.9: Distribution comparison for datasets from three regions, showing significant different activity patterns. Activity type labels are excluded because dataset of Mexico City only contain 10 types of activity that are different from CA and Puget Sound region.

## Data Balancing Algorithm Evaluation

: Data balancing is performed on biased datasets to mitigate issues of skewed representation. To demonstrate its effectiveness, we focused on the most biased dataset in this study: the Mexico City dataset, specifically regarding activity length and activity type. The balancing process reduces the overrepresentation of common activities such as work and school, while enhancing the visibility of less frequent activities. It also reduced the dominance of 3-activity chains, helping to achieve a more even activity distribution. Importantly, while other features were adjusted, the overall activity duration distribution remained consistent. For better visualization, the dominant "Home" activity was excluded.

The outcomes of this balancing effort is elaborated in Fig. 5.10. Fig 5.10. (a) and (c) illustrate the adjusted distributions for activity type and length, with a noticeable reduction in overrepresented activities and chains. Fig 5.10. (b) highlights that the activity duration distribution is preserved despite other changes. Fig 5.10. (d) depicts improvements in model performance metrics, showing significant reductions in JSD for activity type, from 0.038 to 0.009, indicating over 76.3% improvement. Frobenius Norm is also improved by 59.4%, from 0.834 to 0.339. Though temporal metrics shows smaller enhancements due to the focus on activity duration, overall model performance benefits significantly from data balancing, demonstrating its effectiveness in mitigating dataset biases.

### 5.6.4 Validation of ALA and Network Traffic Loading

To validate the results of the ALA method and assess the transportation system-level performance of integrating the Deep Activity model with ALA in a large transportation network, we conduct a series of experiments. As shown in Fig. 5.11 (a), the LA County region is divided into eight sub-regions. We use 100,000 agents from the SCAG ABM dataset as a training set to transfer the mobility model initially trained on the NHTS dataset, and fine-tune the ALA to obtain reference distributions of distances and angles for each sub-region.

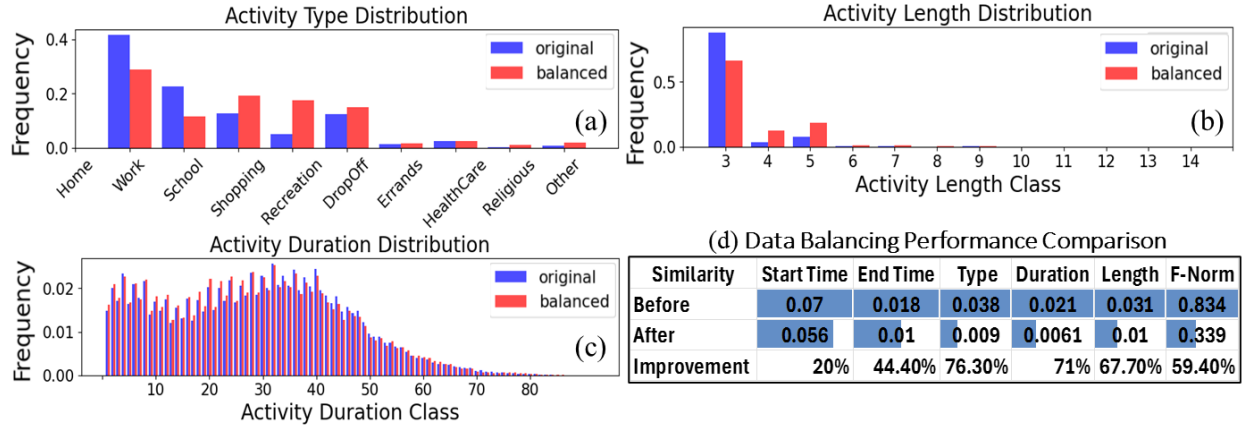


Figure 5.10: Data balancing is performed on Mexico City.

The fine-tuned mobility generation model and ALA are then applied to a population sample of 1 million to evaluate its validity and scalability.

Fig. 5.12 (a) presents the distribution of activity locations across sub-regions in LA County. The results demonstrate that the ALA effectively captures the spatial distribution of activity locations across all sub-regions. The cosine similarity of OD matrices between SCAG ABM and ALA is 0.997, indicating the flow pattern generated by ALA closely aligns with that of SCAG ABM. With the mobility patterns and assigned locations generated, we further load the travel demand into the LA transportation network, as depicted in Fig. 5.11 (b). The system-level traffic performance is illustrated in Fig. 5.12 (b), which shows the hourly VMT and traffic speed over 24 hours, aggregated across all freeway segments. From a network-wide perspective, the proposed Deep Activity model and ALA collectively ensure a well-aligned temporal distribution of traffic flow, yielding MAPEs of 4.97 for VMT and 1.16 for traffic speed when compared to benchmark results from SCAG ABM, demonstrating strong system-level performance.

Beyond the system-wide traffic metrics, we conduct further analysis at the corridor level by selecting a major segment from Interstate 405, a key freeway in the LA network. The location of the selected freeway segment is highlighted by the purple rectangle in Fig. 5.12





(a) Los Angeles County and 8 Sub-Regions



(b) Los Angeles County Freeway Network



(c) MATSim Simulation Network of LA County

Figure 5.11: LA County Map and Freeway Network.



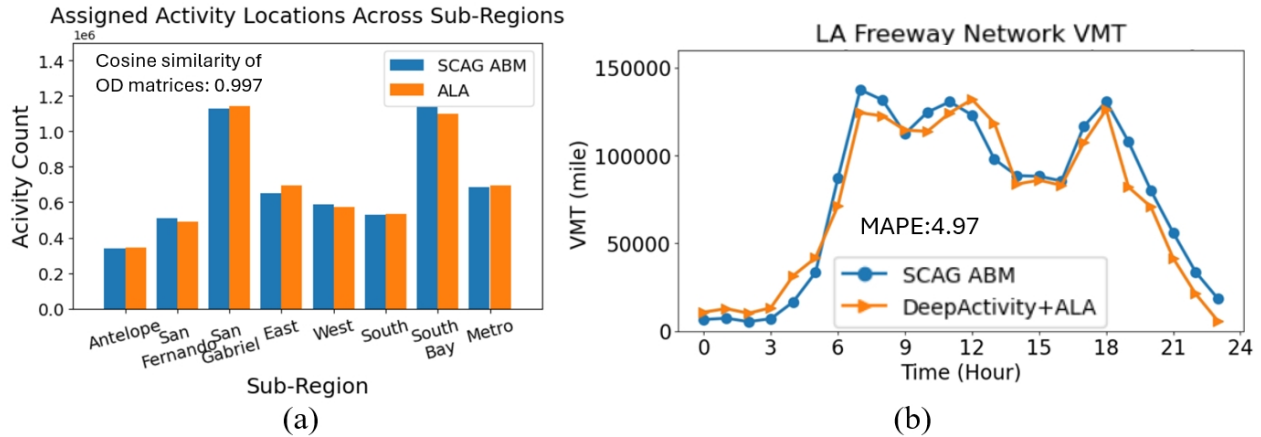


Figure 5.12: Validation for ALA and traffic loading at network level.

(b). The 24-hour traffic volume and speed in both directions on the target corridor segment are shown in Fig. 5.13 for comparison. To ensure the fidelity of the simulation, we include real-world observation data from the Caltrans Performance Measurement System (PeMS) [37] as a reference and compare the results from our proposed model with those from the SCAG ABM data. Note that the MAPE is just calculated between the proposed model results and SCAG ABM, as PeMS data serve only as a reference. As seen in Fig. 5.13, PeMS observations indicate that both directions of the selected corridor experience high traffic volumes and significant congestion during the daytime. Notably, congestion patterns differ between the two directions: the northbound direction experiences major congestion throughout most of the midday, from 6 AM to 8 PM, while the southbound direction’s primary congestion occurs after 12 PM and continues into the evening. The MAPE for traffic volume in the northbound and southbound directions is 5.85 and 9.32, respectively, while the MAPE for traffic speed in the northbound and southbound directions is 4.45 and 4.36, respectively.

These comparisons suggest that the proposed Deep Activity model and ALA successfully capture the dynamic temporal variations in traffic flow at the corridor level, demonstrating the model’s good representation of the transportation system from both a travel demand generation and traffic loading perspective.

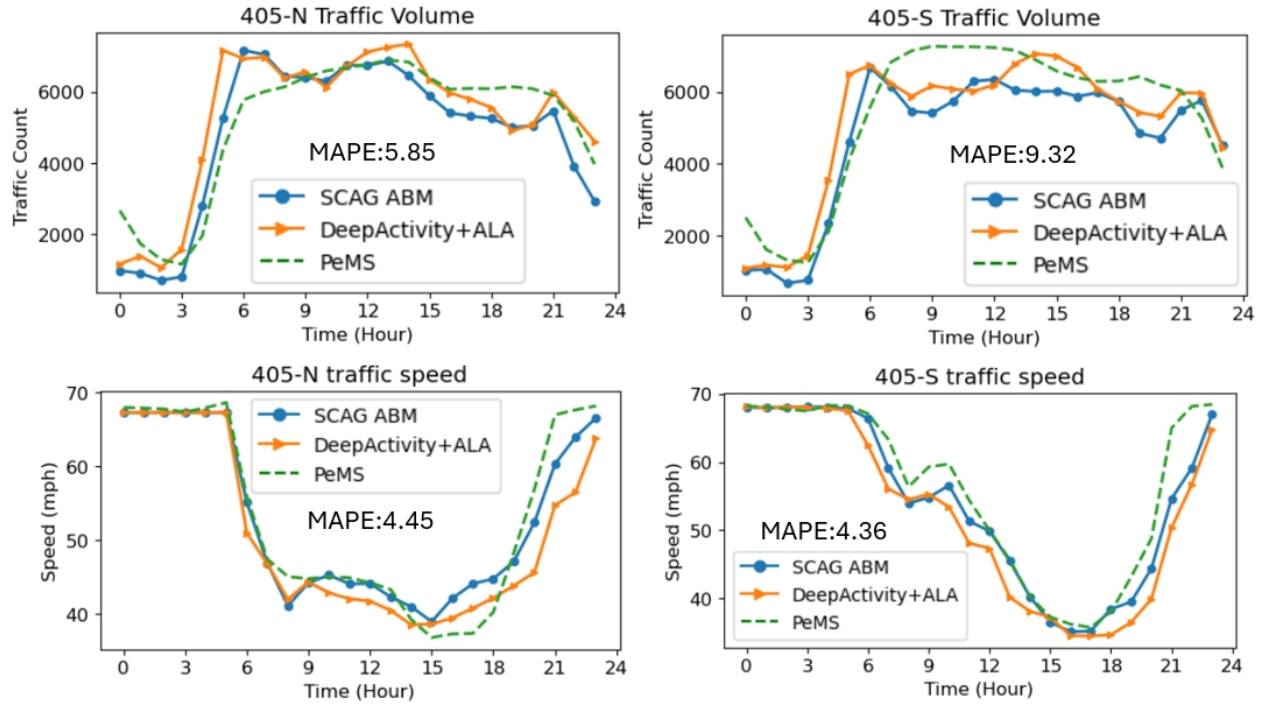


Figure 5.13: Validation for ALA and traffic loading at corridor level.

### 5.6.5 The Influence of Model Complexity and Data Size

The Deep Activity model was trained on the NHTS dataset, which has a modest sample size, potentially limiting the training effectiveness of complex models like Transformers. In this section, we explore the relationship between model complexity and data size in activity generation tasks.

To determine the optimal balance between model complexity and dataset size, we evaluated nine Transformer configurations on the NHTS dataset (180,000 samples). These models, which varied in the number of decoder layers (D), encoder layers (E), and attention heads (H), represented a spectrum of complexity, with parameters ranging from 434,000 to 2,344,000.

Our analysis, as illustrated in Fig. 5.14, indicated that models with fewer layers generally performed better, as reflected by lower JSD values and Frobenius norms. Simpler decoder-only models and those with a balanced encoder-decoder structure showed competitive results.

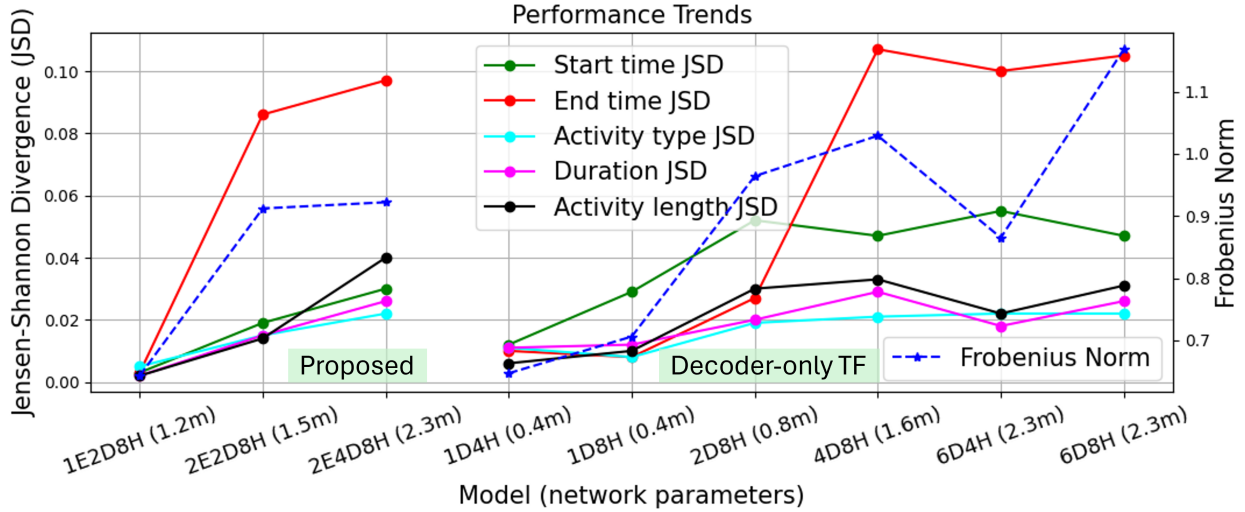


Figure 5.14: Performance evaluation for transformer models with different complexity.

Moreover, increasing the number of attention heads had only a moderate effect on performance. This suggests that, for the given dataset size, adding more layers or attention heads does not guarantee improved performance and may lead to diminishing returns or increased complexity without significant accuracy gains.

To further understand the scaling effects of data size on different models, we compared a Transformer model with an LSTM model, as shown in Fig. 5.15. The Transformer exhibited improved performance with increased data size, from 45,000 to 180,000 samples, whereas the LSTM model’s performance plateaued, indicating its limited ability to benefit from larger datasets. These results highlight the Transformer’s superior capacity for utilizing larger datasets, aided by its parallel processing capabilities and global receptive field.

**Model Selection.** Given the modest sample size of the NHTS dataset, our findings suggest that: **1)** training on NHTS dataset, Transformer-based models with fewer layers perform better; **2)** the number of attention heads has moderate influence on the performance; **3)** Transformer-based model leverage larger datasets better, with performance increased when data size increases; **4)** LSTM-based models reach their limitation in current dataset and plateau with increased data size. Moreover, the explainability (as in Section 5.6.3) and

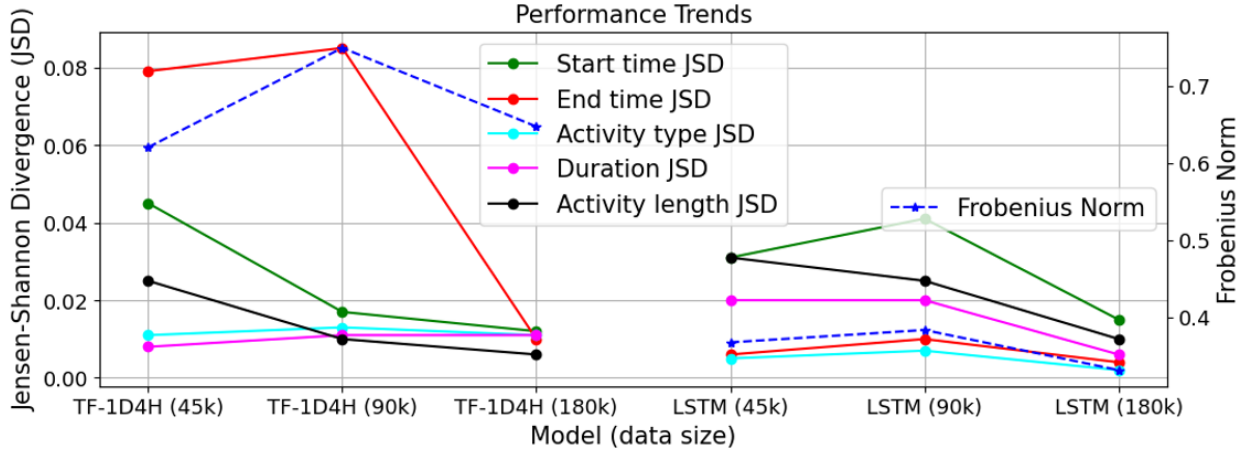


Figure 5.15: Training data size effect on decoder-only transformer and LSTM.

transferability (as in Section 5.6.3) of transformer-based models resulted in choosing the transformer-based Deep Activity model.

## 5.7 Conclusion and Future Work

In this study, we proposed the Deep Activity model, a generative deep learning approach for human mobility modeling. We adopt the concept of "activity chains" to accurately represent the daily mobility patterns of individuals by applying household travel survey data in deep learning to model human mobility patterns, showcasing a pioneering method in the field. The Deep Activity model demonstrated its ability to generate realistic activity chains with high fidelity. Our experiments confirmed the model's robustness and adaptability, showing strong performance when fine-tuned with data from diverse regions such as California, the Puget Sound region, and Mexico City.

**Limitations and Broader Impacts.** While we have established the Deep Activity model, there are opportunities for further exploration. Our current approach is limited by the existing Household Travel Survey (HTS) datasets, which lacks accurate location data. Integrating precise location data into the network could enhance the model's accuracy and

applicability. However, current GPS datasets lack semantic information (e.g., activity type). Therefore, human trajectory data mining to link semantic information with GPS data and create new datasets is crucial [154]. Further integration of location data, activity data, and social-demographic data would facilitate more comprehensive modeling and provide deeper insights into human mobility patterns, supporting advanced applications in urban planning and transportation management.

## Chapter 6

# Deploying scalable traffic prediction models in real-world large transportation networks during hurricane evacuations

Traffic prediction plays a crucial role in facilitating traffic management during hurricane evacuation. Precisely predicting traffic states enables safe and smooth deployment of resources and effective operational strategies by traffic management agencies. However, accurate traffic prediction during evacuation remains challenging due to heterogeneous human behaviors, scarcity of traffic data during evacuation, and uncertainty of hurricane events. This study presents a comprehensive modeling framework that leverages Multilayer Perceptron (MLP) and Long-Short Term Memory (LSTM) models to capture long-term congestion patterns and short-term speed patterns during hurricane evacuation. This framework takes into account various input variables, including archived traffic data, spatial-temporal information of the road network, and hurricane forecast data. The model has been deployed in a real-world traffic prediction system for predicting traffic speed in Louisiana during hurricane evacuation conditions. An evaluation was conducted using archived traffic data from 5 historical hurricanes. The MLP model achieved an accuracy of approximately 82% in

predicting long-term congestion states over a 6-hour period within a 7-day hurricane-impacted duration. Additionally, the short-term speed prediction model demonstrated prediction Mean Absolute Percentage Errors (MAPEs) ranging from 7% to 13% for various evacuation horizons, spanning from 1 hour to 6 hours. The evaluation results demonstrate the model’s potential in enhancing traffic management during hurricane evacuation events. Moreover, the real-world deployment showcases the model’s adaptability and scalability for traffic prediction in diverse hurricane situations within large transportation networks.

## 6.1 Introduction

Over the past few years, several coastal regions in the United States have encountered devastating hurricanes, leading to substantial property damage and loss of lives [138, 288]. These catastrophic events have triggered a renewed focus on the enhancement of evacuation management systems. The effectiveness of evacuations crucially depends on the guidance of evacuation routes and the management of traffic flow [96]. The accuracy of traffic congestion forecasts before hurricanes and real-time traffic state predictions during hurricanes play a pivotal role in achieving these objectives. Providing reliable traffic predictions allows individuals to make well-informed decisions about evacuating, while also enabling emergency management authorities to assess the necessity of issuing evacuation orders.

During recent hurricanes (e.g., Ida 2021), for instance, massive evacuations took place in the southern Louisiana region, especially in its coastal parishes [203]. Severe congestion occurred on several major evacuation routes due to mandatory evacuation orders impacting hundreds of thousands of people [90]. To mitigate the impact of congestion, traffic management agencies can implement various strategies, such as staged evacuation, hard shoulder running, route guidance, and more. However, the success of these measures depends on the accuracy of traffic prediction during the evacuation.

Modeling and predicting traffic conditions during hurricane evacuations presents three

significant challenges. Firstly, the inherent uncertainty of hurricane events encompasses variations in intensity, landfall locations, duration, and more [29]. All these uncertainties lead to changes in the evacuation scale, evacuation direction, evacuation route, size of the evacuating population, and so on, subsequently leading to diverse traffic patterns. Secondly, the decision-making process of evacuees during such critical events is intricate [60, 80, 228], including when to start evacuating, where to evacuate to, which evacuation route to choose, and more. Unlike daily traffic, where commuters follow relatively predictable spatial-temporal patterns, mass evacuations can yield entirely different traffic patterns. Thirdly, the scarcity of high-resolution traffic observation data during historical hurricanes compounds the challenge [6, 222]. In specific regions, powerful hurricanes can be infrequent, resulting in limited available traffic data for evacuation analysis. Take the State of Louisiana as an example, there were only 4 major hurricanes (Category 3 or higher) making landfall on its coast from 2010 to 2023. Moreover, hurricane damage may render road sensors inoperative during evacuations, further constraining the dataset’s usefulness in studying spatial-temporal traffic patterns [247].

Regarding traffic prediction for evacuation scenarios, many previous papers utilized travel demand modeling and traffic assignment simulations to represent the travel behaviors and results of traffic flows during evacuations [148, 179, 223]. These approaches follow typical traffic model principles, using either agent-based or four-step modeling methods to reflect the travel demand generated by evacuation events and simulate the traffic in the transportation network [166, 267]. However, these models require a large amount of survey data to capture people’s activity preferences and develop a series of choice models to generate travel demand, which usually requires calibrating numerous parameters for each sub-model. The nature of this type of modeling makes them a more appropriate tool to analyze existing hurricanes, but they are unable to respond to the rapid evolution of upcoming hurricanes with variable trajectory and intensities, due to the large amount of time and computing resources required to develop these models.

On the contrary, data-driven models, which rely on massive historical data instead of



physical models, have drawn more attention in recent years as an important approach to studying evacuation traffic [9, 26, 200]. With the growing number of detectors installed on roadways, data-driven approaches are more often being leveraged to model traffic flows [94]. With data-driven models, the traffic patterns hidden in historical traffic data can be learned and used to predict the traffic in the future. However, due to the scarcity of hurricane events, the data-driven hurricane evacuation traffic analysis is still not very abundant. Existing data-driven models also have various limitations, such as focusing only on one hurricane [221], which cannot be used to predict hurricanes with variable properties; focusing on one or a few segments of the roadway network [229], which cannot provide predictions for the large-scale network; and requiring explicit certain types of data to run the model training, which cannot be transferred to other regions with only a limited source of historical data.

In 2015, the Federal Highway Administration (FHWA) began developing the Integrated Modeling for Road Condition Prediction (IMRCP) system, a tool that fuses real-time and archived data with results from an ensemble of forecast and probabilistic models [6, 7]. One of the system’s objectives is to enhance its capabilities in traffic predictions during hurricane/tropical storm season for states along the southern coast such as Louisiana, Mississippi, and Alabama. The model proposed in this study, as the core traffic prediction module in the IMRCP system, aims to address three gaps in the field of hurricane evacuation traffic prediction. First, while both long-term planning and short-term response are essential in the planning and management of hurricane evacuation events [19], current emergency traffic prediction models predominantly focus on short-term, real-time predictions due to limitations in the scope of model design. Secondly, hurricane evacuations entail the movement of a significant population across extensive roadway networks, often spanning different cities [19, 72]. However, most existing evacuation traffic models concentrate on specific key corridors or selected road segments, overlooking the broader spatial-temporal correlations across the entire network and their impact on traffic patterns. Lastly, many of the prevailing approaches exhibit a high dependence on data quality and the richness of traffic data features [221, 222, 229].

This focus on comprehensive data requirements often neglects scenarios where a complete dataset is frequently absent in many hurricane-impacted regions, such as cases where only traffic speed data is available.

To address these gaps, this study proposes a multi-scale modeling framework utilizing both Multi-Layer Perceptron (MLP) and Long Short-Term Memory (LSTM) models. The objective is to offer a full-spectrum prediction that encompasses both long-term congestion patterns and short-term speed value predictions for a large-scale transportation network. The primary contributions of this study are summarized as follows:

- **Proposing an integrated hurricane evacuation traffic prediction framework:** This study developed an integrated hurricane evacuation traffic prediction pipeline that encompasses the entire forecast lifecycle, including data acquisition, preprocessing, model training, and deployment. This framework has demonstrated its applicability in a real-world traffic prediction system, IMRCP, and has been proven to provide support for hurricane evacuation management and planning.
- **Facilitating multi-scale and network-wide evacuation traffic predictions:** Utilizing a link-based modeling approach augmented with spatial attributes and dynamic hurricane-related features, our model effectively captures intricate spatial-temporal dependencies between hurricane dynamics and traffic patterns within a comprehensive transportation network. This capability empowers the model to make predictions regarding both long-term congestion states and short-term traffic speeds across an entire state-level transportation network in the context of hurricane evacuation.
- **Applicability to sparse datasets:** Our model introduces a specialized feature engineering and data balancing strategy tailored for hurricane evacuation scenarios, effectively addressing historical data with limited traffic attributes and a scarcity of hurricane records. This strategic approach results in robust prediction performance, facilitating the model's applicability in scenarios characterized by the absence of

comprehensive traffic or hurricane data.

## 6.2 Related Work

### 6.2.1 Model-based evacuation traffic modeling

The majority of model-based traffic evacuation analyses primarily focus on studying the evacuation decisions of evacuees from a behavioral perspective [223]- [148]. Key aspects influencing evacuation travel behaviors include evacuation decisions [270]- [98], estimated travel time [97], departure time [233], and destination choices [224]. While numerous studies have explored the impact of individual decision-making processes during evacuation [166] [267] on resulting travel demand, these approaches mainly rely on traditional methods, such as collecting survey data, which may prove inadequate for real-time hurricane evacuation scenarios. Lindell et al. [148] summarized the modeling procedures and components of large-scale evacuation processes; however, only a few models have been validated, and they are validated mainly at regional scales [66] [176] [279]. Among model-driven approaches, only a limited number of model-based methods have addressed traffic patterns during evacuation, primarily centered on analyzing highway capacity loss during the evacuation process [61] [244]. Moreover, the process of gathering and calibrating parameters for model-based approaches often proves laborious and time-consuming, leading to a growing inclination towards utilizing data-driven methods as a viable alternative.

### 6.2.2 Data-driven evacuation traffic modeling

Data-driven methods enable the prediction of future traffic states by analyzing historical traffic data to discern traffic patterns. Several factors contribute to the growing popularity of data-driven approaches: firstly, model-based approaches demand tedious, labor-intensive, and time-consuming calibration efforts [86]. Secondly, the widespread deployment of diverse traffic

sensors on roadway networks facilitates the use of methods reliant on extensive archived traffic data [200] [84] [258]. Thirdly, data-driven methods offer streamlined and generalized model training and updating frameworks, easing their transferability to different regions with reduced additional model development and maintenance efforts [26]. These data-driven approaches find applications in various scenarios, such as speed prediction [47], travel time prediction [9], and traffic flow prediction [94].

However, early-stage data-driven methods predominantly rely on simplistic machine learning models such as Support Vector Machine (SVM) [281], K-Nearest Neighbor (KNN) [283], and Artificial Neural Network (ANN) [211]. Their limitations lie in handling large-scale networks and complex traffic patterns. As a consequence, an upward trend involves employing deep neural networks for traffic prediction, including Recurrent Neural Networks (RNN) [180] [236] [248], Convolutional Neural Networks (CNN) [221] [229], Graph Convolutional Neural Networks (GCN) [8] [68], and their combined forms [6] [42]. Nevertheless, in the context of hurricane evacuation scenarios, the scarcity of specific traffic data during hurricane periods poses additional challenges in adopting the aforementioned models, which demand extensive data for robust performance.

Although the above methods have been used in the general traffic prediction field, to the best of our knowledge, only a few papers have touched on using deep learning models to predict traffic during hurricane evacuation [223] [221] [229]. The problem with these studies is that they focus only on one small segment from the freeway network and only use one historical hurricane for training, validation, and testing, which lack scalability and transferability. The only paper that has studied the network-level traffic prediction issue is presented by Rahman et al. [222]. This model first introduced network-wise traffic prediction using a graph convolution LSTM model. However, the author only used Hurricane Irma's historical traffic data for model development and testing, therefore the model's performance is unknown if implemented in other hurricanes.

## 6.3 Methods

### 6.3.1 Methodology Overview

In general traffic predictions, the prediction can be divided into short-term and long-term forecasts [263]. A time span of fewer than one hour is usually considered a short-term forecast, otherwise, it is considered a medium and long-term forecast [109] [282]. In the context of emergency evacuation, however, the requirement for long-term traffic forecasts extends far beyond one hour into the future. Due to the nature of hurricane events where hurricanes continuously change intensity and position, the planning for hurricane evacuation events requires two-fold support from the traffic prediction model [19]: 1) long-term traffic information that can be generated well ahead of the hurricane (e.g., multiple days) to assist in advance planning and coordination of potential support, and 2) short-term real-time prediction to provide dynamic changes in the traffic pattern (e.g., a few hours) for quick emergency response.

Based on these two requirements, our multi-scale hurricane evacuation traffic prediction model proposes two modules that tackle different aspects of the evacuation planning requirements: 1) a long-term prediction module, which provides a prediction with a long span into the future with low time granularity, this helps initiate the evacuation plan days before the hurricane makes landfall, and 2) a short-term model, which focuses on the near future but with high time granularity, this helps evacuation agencies perform timely responsive operations to the most updated road traffic conditions based on the real-time traffic data fed.

Fig 6.1 illustrates the comprehensive framework of the hurricane evacuation traffic prediction model proposed in this study. This model utilizes three main categories of input data: traffic data, hurricane data, and network geospatial data. This comprehensive set of input features enables the model to learn the interdependencies between various types of hurricane events and the corresponding traffic patterns during evacuations. The input data encompasses both historical data from past hurricane evacuation events for training purposes and real-time

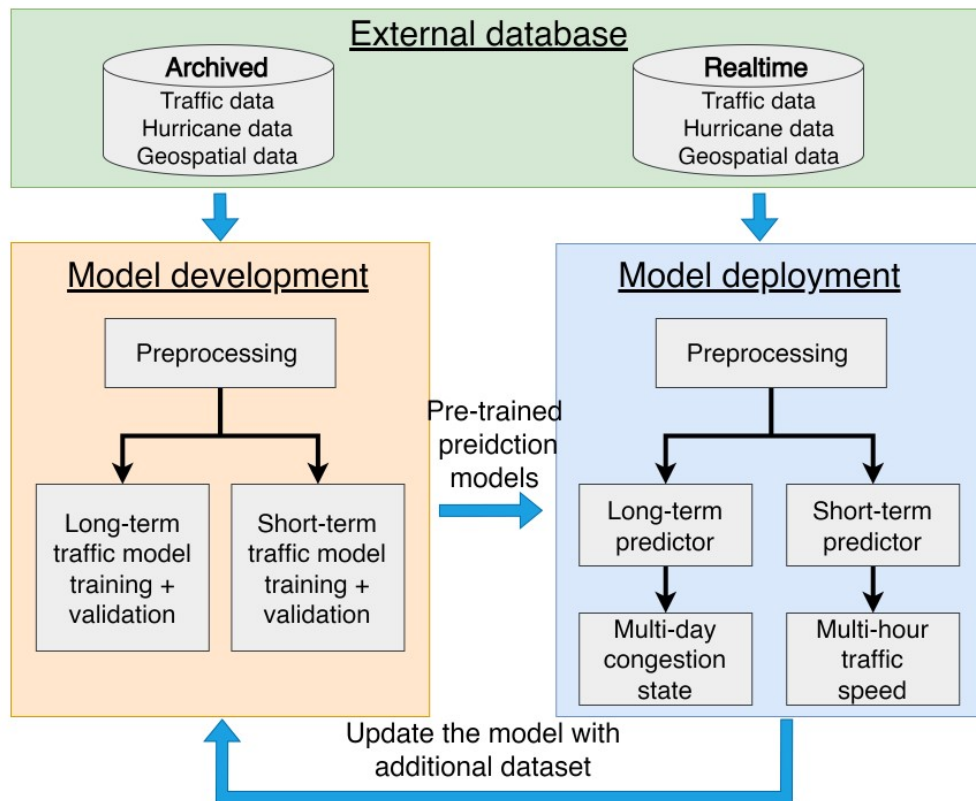


Figure 6.1: The framework of the multi-scale hurricane evacuation traffic prediction model

data during active hurricane events for deployment. The model proposed in this study, being one of the core traffic prediction components within the IMRCP system [7], utilizes IMRCP’s integrated data sources for both development and deployment phases. During the model development phase, multi-source historical data from previous hurricane events are extracted from the IMRCP system and leveraged to train both long-term and short-term prediction models. In the deployment phase, the model utilizes the pre-trained prediction model as a predictor and leverages real-time data from IMRCP to provide predictions for both the long-term congestion state and short-term traffic speed. After each new hurricane event, the data is archived and subsequently transferred to the model development module for the purpose of updating the prediction models.

### 6.3.2 Long-term Congestion State Prediction Model

The long-term model’s primary objective is to offer multi-day predictions concerning the location and timing of congestion during a hurricane. Many existing multi-day traffic analysis studies use a 7-day or one-week duration as the minimum duration for traffic pattern analysis [26] [221]. Therefore, we also adopt a 7-day span as the prediction duration for long-term congestion prediction during hurricane evacuation. To encompass traffic patterns both before and after the hurricane, the 7-day prediction horizon is defined as follows: 3 days before the landfall, the day of landfall, and 3 days after the landfall. Given the extended prediction duration, the long-term model emphasizes the network-level spatial-temporal distribution of potential congestion at a sparse time resolution (e.g., 6 hours over 7 days), rather than providing precise speed value predictions at a high temporal resolution.

In contrast to short-term traffic prediction, where real-time traffic data continuously feeds into the model, the long-term prediction model needs to provide 7-day predictions without any online data input during this period. This technical challenge renders regression prediction impractical. Consequently, we transformed this problem into a multi-class classification task,

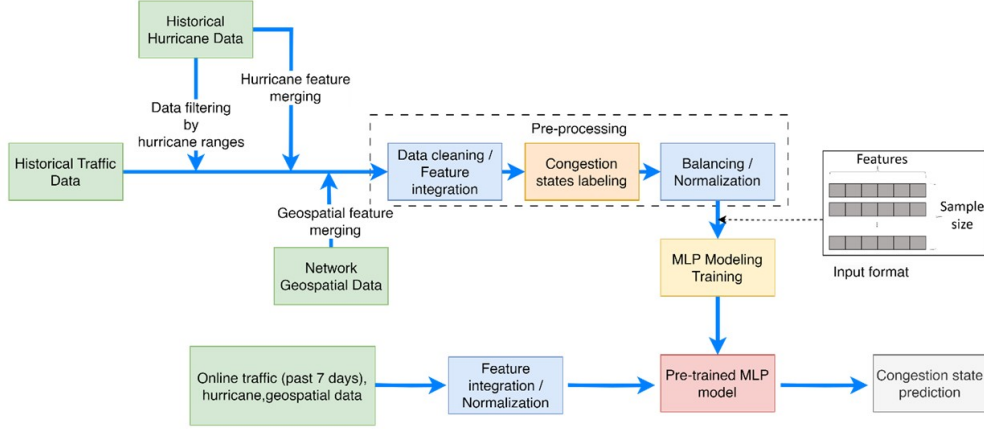


Figure 6.2: Model framework for long-term hurricane evacuation traffic prediction model

predicting the congestion state for each sub-segment (e.g., 6-hour) across the 7-day duration.

Assume the input feature for a single time slot can be represented by:

$$S_i = \{x_h, x_t, x_{sp}, x_l\} \quad (6.1)$$

The output can be expressed as:

$$C_i = MC(X_i) \quad (6.2)$$

In this study, we adopted the MLP as the classification model for the long-term congestion state prediction. The MLP model proves particularly suitable for complex and non-linear classification tasks with intricate relationships between input features and the target class. The model's hidden layers enable it to learn intricate patterns and representations from the data, effectively handling complex decision boundaries and capturing high-dimensional feature interactions. Compared to traditional classification models like Logistic Regression, SVM, or Decision Trees, the MLP exhibits greater power in modeling complex relationships and achieving higher accuracy [223] [86] [200]. To prepare the training samples, we conduct data aggregation and annotation on original high-resolution (e.g., 5-minute interval) traffic data from historical hurricanes, identifying congestion states for each 6-hour period across



the 7-day horizon. For determining the congestion state, we adopt the speed performance index (SPI) from previous traffic congestion studies [68], representing the ratio of average speed to the maximum permissible speed:

$$SPI = \frac{v_{\text{mean}}}{v_p} \times 100\% \quad (6.3)$$

In this study, we use the weekly average speed from the previous week to represent the maximum permissible speed. Congestion states are categorized as heavy congestion ( $SPI < 50\%$ ), light congestion ( $50\% < SPI < 75\%$ ), and no congestion ( $SPI > 75\%$ ).

Fig 6.2 illustrates the model framework for the long-term congestion state prediction module. This model receives various sources of data as input, including historical traffic data, historical hurricane data, and network geospatial data. Upon integrating these diverse features, the samples are labeled using the criteria outlined in Equation 6.3. Subsequently, we conduct data balancing based on congestion labels to ensure the dataset’s impartiality with respect to different classes. Following this, we normalize the input data and feed it into the core training module. For the multi-class classification task in this study, we employ an MLP neural network, comprising an input layer, multiple hidden layers, and an output layer.

Firstly, the input layer serves as the initial processing stage, receiving and transmitting input data to subsequent layers. Each node (neuron) in this layer corresponds to a specific feature or attribute present in the input data. The input features utilized in our model can be categorized into four distinct groups. These groups encompass link features, such as the number of lanes, directions, and non-evacuation average speed, providing essential transportation link characteristics. Spatial features, including latitude, longitude, and distance to the landfall location, contribute valuable spatial context. Temporal features, such as time of day and time to landfall, offer crucial temporal information. Lastly, the hurricane features, encompassing forecast hurricane category and potential landfall location, provide essential hurricane-related data.

Secondly, the hidden layer plays a pivotal role as intermediary layers responsible for information processing. Each node in a hidden layer receives inputs from all nodes in the preceding layer and forwards its output to all nodes in the subsequent layer. This configuration allows the neural network to learn intricate patterns and relationships present in the data across different categories of features in the input layer.

Lastly, the output layer generates the final predictions or outputs of the model. In our long-term prediction model, the output layer offers predictions among three congestion labels: no congestion, light congestion, and heavy congestion. This final prediction from the output layer reflects the model’s evaluation of the congestion state during the hurricane evacuation process.

### 6.3.3 Short-term Traffic Speed Prediction Model

The short-term model focuses on accurately predicting speed values for each link during the 7-day hurricane impact horizon, as defined in the problem statement. It aims to forecast speeds for a short-term horizon ranging from one to several hours into the future, given a particular start time. The short-term speed prediction is framed as a many-to-one time series regression problem, where the input includes time sequence data (e.g., speed data for the last 24 hours) along with link attributes, spatial-temporal attributes, and the latest hurricane attributes. The output is the predicted speed value after the specified prediction horizon from the current time step.

Assume we have a sequence of data for the last  $N$  time steps:

$$S_i = \{y_1, y_2, \dots, y_N\} \tag{6.4}$$

where each data point  $y_i$  contains the traffic speed, link attributes, hurricane forecast, and spatial-temporal features for a specific link. We want to predict the speed value that is  $k$  time steps from now:

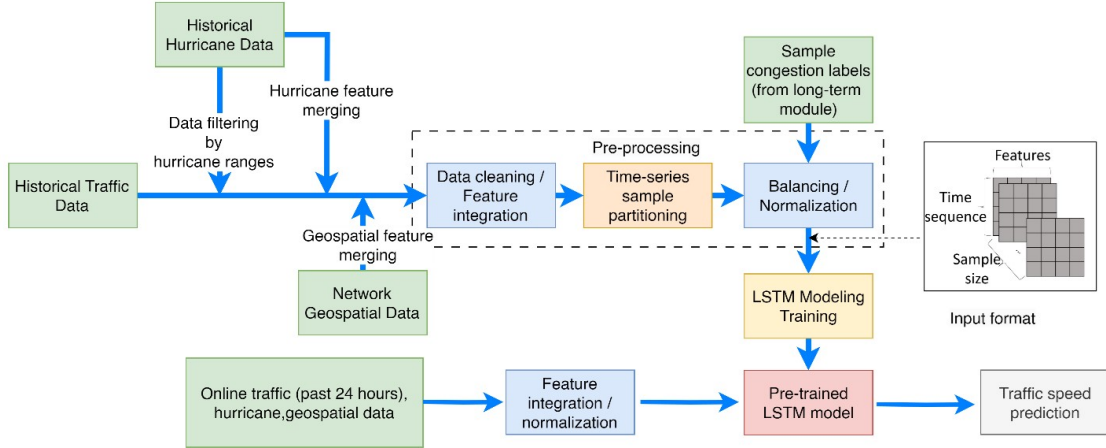


Figure 6.3: Model framework for short-term hurricane evacuation traffic prediction model

$$Y_{N+k} = TS(X) \quad (6.5)$$

where  $k$  is defined as the prediction horizon.

Fig 6.3 illustrates the model framework for the short-term traffic speed prediction module. The short-term module utilizes data resources similar to the long-term module as input. Following the integration of features, a crucial step in the short-term module involves partitioning the raw time-series data into small sequential samples of equal sequence length. To maximize the generation of samples, especially in cases of limited time-series data during historical hurricane evacuation events, we have adopted a sliding-window approach that has previously been employed in other time-series forecasting problems. This approach allows us to extract numerous small sequential data segments from the same road link sample over multiple days. Subsequently, we also perform data balancing on the partitioned samples. In this stage, we utilize pre-generated labeled data from the long-term module to ensure a balanced representation of time-series samples across different congestion states. After balancing and normalizing the pre-processed data, we then proceed to feed the data into the model training procedure.

Given the complex non-linearity and spatial-temporal dependencies between speed pat-

terns, links, and hurricanes, the short-term prediction adopts LSTM, an advanced form of RNN that can overcome the disadvantages of RNNs such as gradient vanishing when dealing with long input sequences. LSTM is specifically designed to handle sequential data, making it highly effective in capturing long-term dependencies [6]. These characteristics render LSTM a suitable choice for real-time traffic prediction during hurricane evacuation scenarios.

The cell state is a crucial component of the LSTM as it serves as a memory that runs through the entire sequence. It allows relevant information to persist over long sequences, enabling the model to capture long-term dependencies effectively. The cell state is updated at each time step using the gates and the previous cell state. In an LSTM, the cell state (hidden state) is divided into two states: short-term state ( $h_t$ ) (similar to an RNN) and long-term state ( $c_t$ ). The long-term state ( $c_t$ ) stores the information to capture the long-term dependencies among the current hidden state and previous hidden states over time. Traversing from the left to the right, the long-term state passes through a forget gate and drops some memories and then adds some new memories via an additional operation. A fully connected LSTM cell contains four layers (sigma and tanh), and the input vector ( $x_t$ ) and the previous short-term state ( $h_{t-1}$ ) are fed into these layers. The main layer uses tanh activation functions which output ( $g_t$ ). The output from this layer is partially stored in the long-term state ( $c_t$ ). The other three layers are gate controllers using logistic activation functions, and their output ranges from 0 to 1. The forget gate controls which parts of the long-term state should be erased, while the input gate in the middle decides which parts of the input should be added. The output gate finally controls which parts of the long-term state should be read and output at this time step ( $y_t$ ).

Throughout the sequential processing, the LSTM iteratively updates its cell state and hidden state, considering the current input and the information from previous time steps. The final hidden state at the last time step can be used for making predictions or passed as input to other layers of the neural network for further processing. This capacity to control the flow of information through the cell state and the presence of the forget, input, and

output gates enable the LSTM model to learn and capture long-term dependencies effectively, making it a powerful tool for various tasks involving sequences.

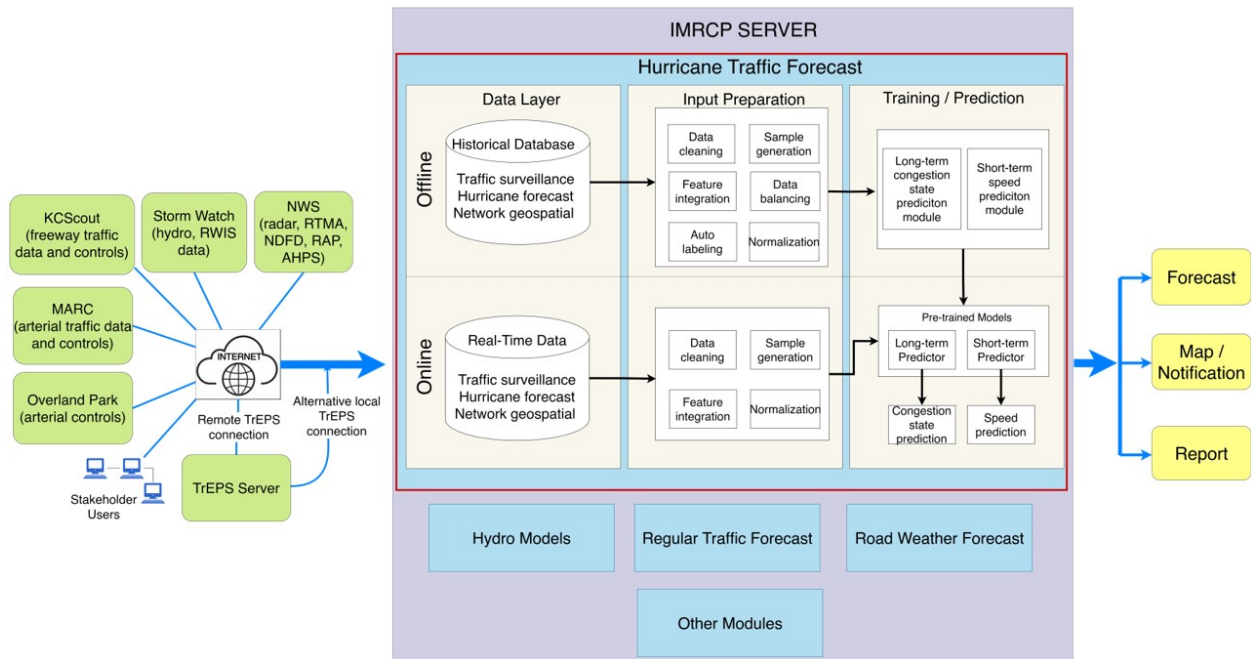


Figure 6.4: IMRCP system integration diagram

## 6.4 Experiment Design

### 6.4.1 Data Preparation

In 2015, the FHWA Road Weather Management Program (RWMP) began developing the IMRCP system to investigate and capture the potential for operational improvements [6]. The resulting IMRCP tool incorporates real-time and archived data from various data sources with results from an ensemble of forecast and statistic models, fusing them together in order to predict the current and future overall road/travel conditions for travelers, transportation operators, and maintenance providers. Besides different types of forecasts, IMRCP also provides flexible reporting tools and an interactive map to meet those needs. The data

that populates these user interface features are kept in a data store that contains collected data and data generated through forecasting components. An illustration of the correlation between each component in the IMRCP system is shown in Fig 6.4.

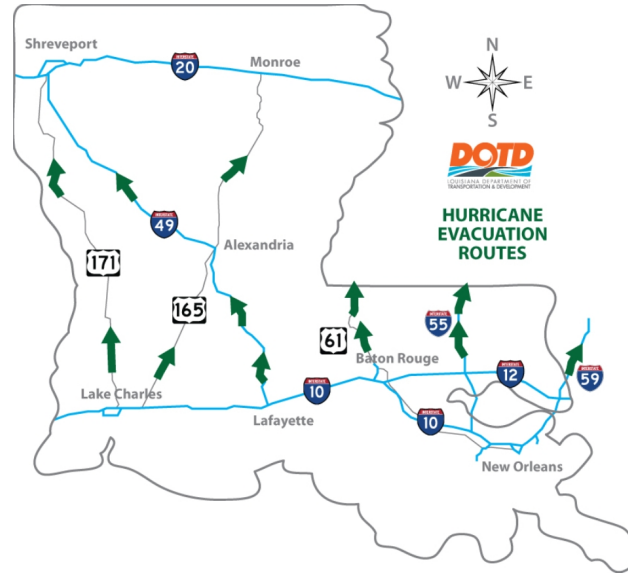


Figure 6.5: Roadway network of Louisiana

The latest phase of IMRCP is engaging multiple state agencies in Louisiana, Mississippi, and Alabama to apply IMRCP capabilities in order to expand situational awareness, planning, and response to extreme weather and operational events, such as hurricane evacuations, as highlighted in the IMRCP Module in Fig 6.4. The Phase 5 IMRCP adopts the traffic speed prediction model proposed in this study as the hurricane evacuation traffic speed prediction module. The model is developed based on historical traffic speed and hurricane data from the past four years in the state of Louisiana. The transportation network of Louisiana is shown in Fig 6.5. There are about 8,000 segmented roadway links along the evacuation routes suggested by LADOTD, which contain all interstate freeways and some U.S. highways in Louisiana.

The total dataset contains five separate datasets corresponding to 5 hurricanes that made landfall in Louisiana from 2019 to 2021, as shown in Table 6.1. Each dataset covers a 7-day time range, including 3 days before landfall, the day of landfall, and 3 days after the landfall.

Table 6.1: Dataset description

Dataset	Hurricane Name	Hurricane Category	Landfall Time	Landfall Location	Traffic Data Source	Hurricane Data Source
1	Ida	4	8/29/2021 16:55	(29.1°N, 90.2°W)	IMRCP	NHC
2	Delta	4	10/9/2020 23:00	(29.8°N, 93.1°W)		
3	Laura	4	8/27/2020 6:00	(29.8°N, 93.3°W)		
4	Zeta	3	10/28/2020 21:00	(29.2°N, 90.6°W)		
5	Barry	1	7/13/2019 15:00	(29.6°N, 92.2°W)		

The speed data is recorded at an interval of 5 minutes, as shown in Fig 6.6. In the example from Fig 6.6, we selected a link on westbound I-10 at (30.19 N, 93.32 W) and plotted its speed across the 7-day hurricane horizon. We can find that the 7-day speed pattern illustrates a clear congestion pattern lasting for over 20 hours about a day prior to the hurricane’s landfall, which could be regarded as evacuation-induced traffic congestion. The hurricane-related data are extracted from the National Hurricane Center (NHC) [203].

Table 6.2 provides the description of variables used for model training. As shown in Table 6.2, the variables are categorized into four groups: link-related, spatial, temporal, and hurricane-related. The only different variables for the long-term and short-term models are the temporal variables. In the long-term model, the time-of-day variable has four classes, each representing a 6-hour period of the day, and the time-to-landfall variable is calculated as the number of 6-hour periods from the current time to the landfall time. On the other hand, for the short-term model, the temporal variables adopt the hour of the day instead of the time of the day, and the time-to-landfall is defined as the number of hours to the landfall time.

## 6.4.2 Baseline Models

To evaluate the significance and benefits of using the MLP and LSTM, we also tested a few other machine learning models as a baseline performance reference. By comparing the results of the proposed models to the baseline models, we can better understand whether the

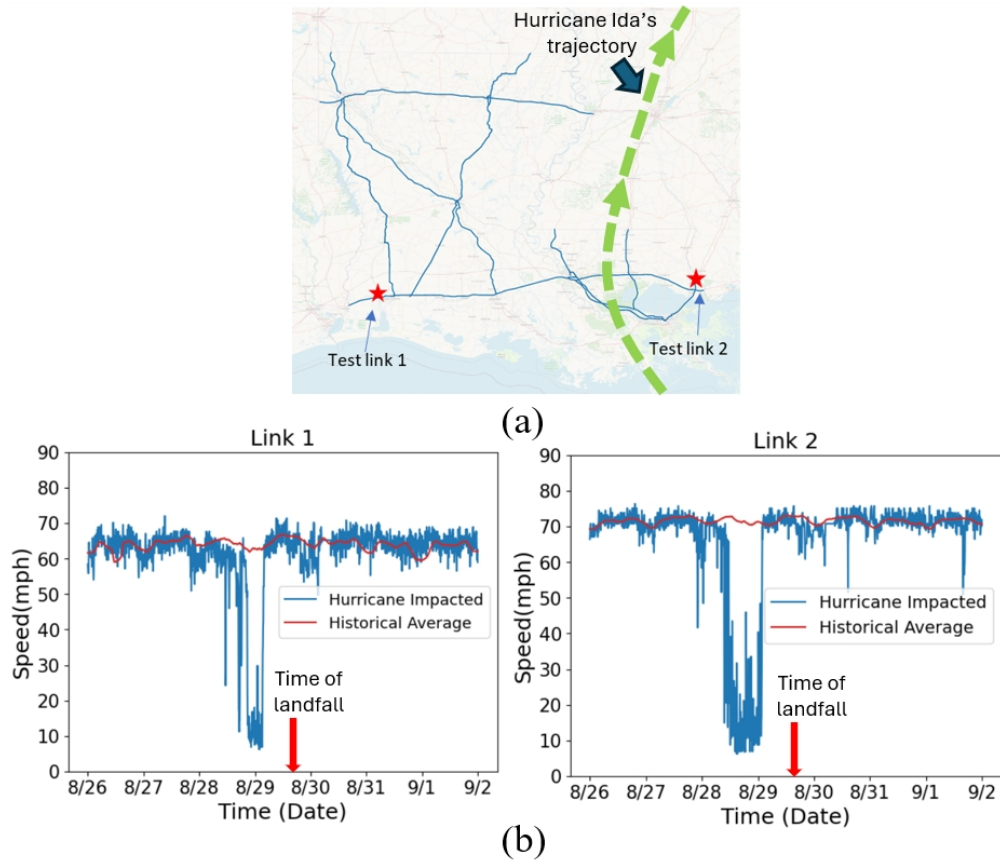


Figure 6.6: Example speed data during hurricane evacuation: (a) Link location; (b) 7-day speed plot of an example link during hurricane Ida



Table 6.2: Input variable description

Variable Group	Variable Name	Description
Link /Traffic	Direction	Direction of the link
	Lanes	Number of lanes of the link
	Regular mean speed	The average speed on this link during the past 7 days.
	Regular speed standard deviation	The standard deviation of the speed on this link during the past 7 days
	Speed (for short-term prediction)	A sequence of traffic speed data for the past 24 hours
Spatial	Latitude	Latitude of the link centroid
	Longitude	Longitude of the link centroid
	Distance to landfall	Distance from the link centroid to the landfall location. (km)
Temporal	Time of day (long-term prediction)	It indicates which 6-hour time slot the current time falls in. 1 indicates 0:00~6:00, 2 indicates 6:00~12:00, 3 indicates 12:00~18:00, and 4 indicates 18:00~24:00
	Time to landfall (long-term prediction)	Time to landfall. It indicates the time difference from landfall. In the long-term model, the time difference is calculated as the number of 6-hour time slots between the current time and landfall time.
	Hour of day (short-term prediction)	Hour of the day for this data record (0,1, 2, ... 22,23)
	Time to landfall (short-term prediction)	Time to landfall. It indicates the time difference from landfall. In the long-term model, the time difference is calculated as the number of 6-hour time slots between the current time and landfall time.
Hurricane	Category	Hurricane category
	Landfall zone	Landfall location. It's a binary variable to indicate on which half of the coastline the hurricane made landfall. 0 indicates west and 1 indicates east
Output	Congestion Label (long-term prediction)	Congestion status for each 6-hour period, 0 for no congestion, 1 for light congestion, 2 for heavy congestion
	Speed (short-term prediction)	Traffic speed 1 to 6 hours from now.

Table 6.3: Summary of model parameters

Model	Parameter	Best parameters
MLP	Number of hidden layers	3
	Number of nodes per layer	64
	Optimizer	adam
	Number of epochs	150
	Batch size	32
	Learning rate	0.001
LSTM	Number of LSTM cells	64
	Number of LSTM layers	2
	Dropout	0.5
	Optimizer	adam
	Learning rate	0.001
	Batch size	16
	Number of epochs	100

additional complexity of the new model is justified by its performance gains.

For long-term congestion pattern prediction, we selected the following models as baseline models:

- **K-nearest neighbors algorithm (KNN).** KNN is a non-parametric, supervised learning classifier, which uses proximity to make classifications or predictions about the grouping of an individual data point. While it can be used for either regression or classification problems, it is typically used as a classification algorithm, working off the assumption that similar points can be found near one another.
- **Support vector machine (SVM).** SVM is a supervised learning model used mainly for binary classification tasks, but it can be extended to multi-class classification as well. The primary objective of SVM is to find the optimal hyperplane that best separates the data points belonging to different classes in a high-dimensional feature space.

For short-term traffic speed prediction, we selected the following two models as baseline models:

- **AutoRegressive Integrated Moving Average (ARIMA).** ARIMA is a time series forecasting model used for analyzing and forecasting time-dependent data. It is a combination of three components: AutoRegressive (AR), Integrated (I), and Moving Average (MA). ARIMA is widely used for time series forecasting when the data exhibits trends, seasonality, and autoregressive patterns.
- **Vanilla RNN model.** Vanilla RNN, also known as Simple RNN, is the basic and original form of a recurrent neural network. It is a type of artificial neural network designed for processing sequential data, such as time series, natural language, and audio data. The basic RNN has a feedback loop that allows information to persist from one time step to the next, making it suitable for sequential data.

For both the long-term and short-term prediction models, we used a stratified 6:2:2 data split for the training, validation, and testing sets, repeating this process five times for cross-validation. In each iteration, the model was trained on the training set and evaluated on the testing set of the current split. The results from these five iterations were then averaged to provide a reliable assessment of model performance. Table 6.3 presents a summary of the estimated parameters for the MLP and LSTM models in our study. We implemented all the models in Python. Unless otherwise specified in Table 6.3, we have used the default parameters of PyTorch for MLP and LSTM.

### 6.4.3 Evaluation Metrics

A comprehensive set of evaluation metrics was chosen to evaluate the prediction performance of both the long-term congestion prediction model and the short-term speed prediction model.

For the long-term speed pattern prediction model, which involves a multi-class classification problem, we utilized elements from the confusion matrix, specifically True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN), to calculate the following performance metrics:

- **Accuracy** is the most straightforward metric and represents the ratio of correctly predicted samples to the total number of samples in the dataset. It measures the overall performance of the model across all classes:

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (6.6)$$

- **Precision**, also known as Positive Predictive Value (PPV), measures the proportion of true positive predictions (correctly predicted positive samples) out of all positive predictions made by the model. It indicates how many of the positive predictions are

actually correct:

$$\text{precision} = \frac{TP}{TP + FP} \quad (6.7)$$

- **Recall** is the proportion of true positive predictions (correctly predicted positive samples) out of all actual positive samples in the dataset. It indicates the model's ability to correctly identify positive samples:

$$\text{recall} = \frac{TP}{TP + FN} \quad (6.8)$$

- **F1 score** is the harmonic mean of precision and recall. It provides a balance between precision and recall and is useful when both false positives and false negatives are equally important. The F1 score ranges from 0 to 1, with 1 being the best possible score:

$$F1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (6.9)$$

In the evaluation of hurricane congestion prediction, recall is more important than precision. This is because ensuring the safety of evacuating residents is the top priority. Predicting congested road links allows authorities to anticipate potential bottlenecks and take proactive measures to prevent evacuees from being exposed to hazardous conditions. Therefore, our model prioritizes whether the majority of actually congested links are predicted as congested, even if this means the results come with a high number of false positives.

For the short-term model, we utilized the following metrics:

- **Root Mean Squared Error (RMSE)** is a measure of the average difference between the predicted values and the actual target values. It is calculated by taking the square root of the mean of the squared differences between the predicted and actual values:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_{\text{true}}^i - y_{\text{predict}}^i)^2}$$

- **Mean Absolute Error (MAE)** is another measure of the average difference between the predicted values and the actual target values. It is calculated by taking the mean of the absolute differences between the predicted and actual values. Unlike RMSE, MAE does not square the errors, so it treats all errors equally regardless of their magnitude:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_{\text{true}}^i - y_{\text{predict}}^i|$$

- **Mean Absolute Percentage Error (MAPE)** is a relative measure of the average difference between the predicted values and the actual target values. It calculates the percentage difference between the predicted and actual values and then takes the mean of these percentage differences:

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|y_{\text{true}}^i - y_{\text{predict}}^i|}{y_{\text{true}}^i} \times 100$$

## 6.5 Results

### 6.5.1 Long-term congestion state prediction

Table 6.4 presents the prediction performance of various models in forecasting long-term congestion patterns during hurricane evacuation. Our proposed MLP model stands out, outperforming all baseline models across all three congestion labels and performance metrics. This outcome highlights the efficacy of the MLP architecture, with its multiple hidden layers and activation functions, in effectively handling complex and non-linear relationships between input features and target classes, particularly in the context of hurricane congestion patterns.

The MLP model achieves an impressive total accuracy of 82%, indicating its ability to accurately predict the locations and timings of congestion during hurricane evacuation periods. Notably, the prediction performance for heavy congestion surpasses that of light

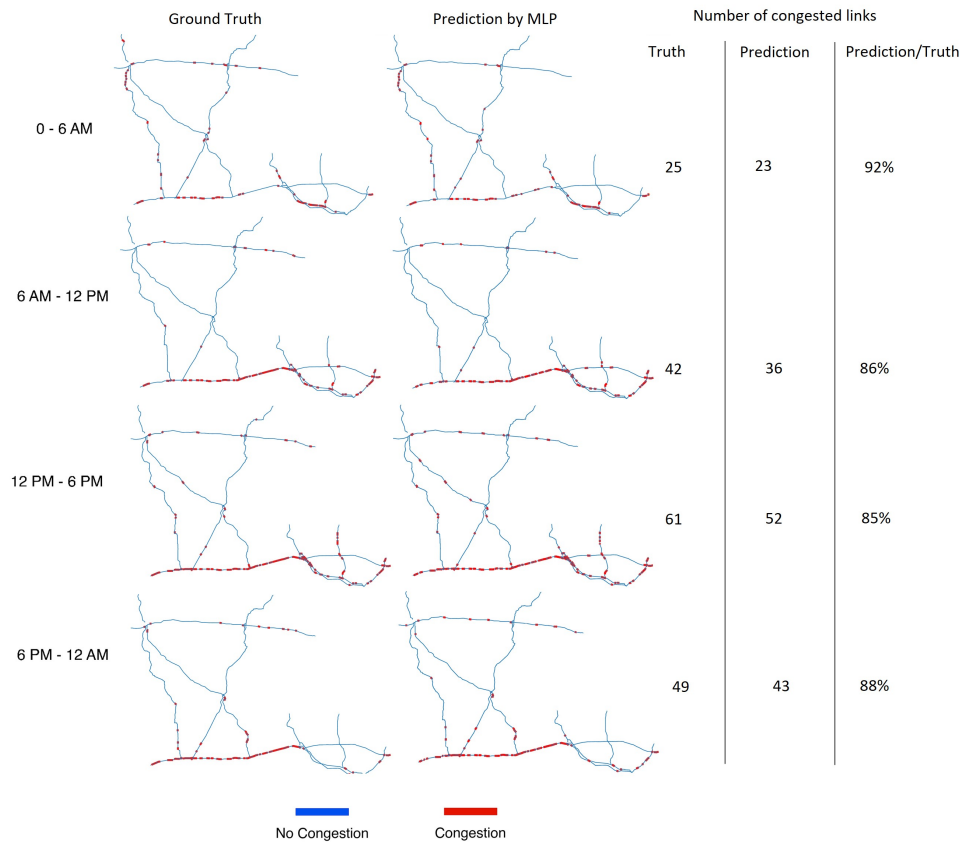


Figure 6.7: Example speed data during hurricane evacuation: (a) Link location; (b) 7-day speed plot of an example link during hurricane Ida

congestion in all three models (KNN, SVM, and MLP). The precision, recall, and F1 scores are consistently higher for heavy congestion compared to light congestion. This discrepancy can be attributed to the nature of heavy congestion, which is more likely to occur in specific locations, such as freeways outbound from cities, and during specific time periods, like one or two days before hurricane landfall. As a result, its spatial-temporal pattern exhibits a stronger correlation with temporal events, such as hurricane evacuations. In contrast, light congestion may occur due to various localized incidents and may not necessarily correlate directly with hurricane evacuation. Consequently, the spatial and temporal patterns for light congestion are less significant compared to heavy congestion.

Table 6.4: Performance of MLP and other baseline models on long-term congestion prediction (average  $\pm$  standard deviation across 5 experimental repeats)

Model	Label	Metrics			
		Precision	Recall	F1 score	Accuracy
KNN	No Congestion	$0.78 \pm 0.011$	$0.72 \pm 0.021$	$0.75 \pm 0.022$	$0.75 \pm 0.019$
	Light Congestion	$0.72 \pm 0.011$	$0.76 \pm 0.014$	$0.74 \pm 0.021$	
	Heavy Congestion	$0.81 \pm 0.034$	$0.77 \pm 0.012$	$0.79 \pm 0.014$	
SVM	No Congestion	$0.81 \pm 0.025$	$0.80 \pm 0.033$	$0.80 \pm 0.039$	$0.79 \pm 0.032$
	Light Congestion	$0.74 \pm 0.009$	$0.76 \pm 0.009$	$0.75 \pm 0.013$	
	Heavy Congestion	$0.85 \pm 0.012$	$0.78 \pm 0.004$	$0.81 \pm 0.008$	
MLP	No Congestion	$0.81 \pm 0.021$	<b><math>0.84 \pm 0.027</math></b>	$0.83 \pm 0.017$	<b><math>0.82 \pm 0.021</math></b>
	Light Congestion	$0.76 \pm 0.008$	<b><math>0.83 \pm 0.006</math></b>	$0.79 \pm 0.018$	
	Heavy Congestion	$0.85 \pm 0.011$	<b><math>0.86 \pm 0.011</math></b>	$0.86 \pm 0.025$	

It is worth noting that while the MLP model only increases the overall prediction accuracy from 79% (SVM) to 82%, its impact on recall for congestion is quite significant. The recall for light congestion increases from 76% to 83%, and for heavy congestion, it improves from 78% to 86%. These enhancements indicate that despite the MLP introducing more false positives, the prediction performance for real-congested links has been greatly improved. This highlights the importance of adopting the MLP model in traffic congestion prediction during hurricane evacuation.

To visually demonstrate the spatial distribution of hurricane-induced congestion during evacuation, we focused on a specific day (one day before hurricane Ida’s landfall) to evaluate

the performance of our long-term congestion pattern prediction model. Fig 6.7 illustrates a comparison between the ground truth and prediction results generated by the MLP model for four distinct time periods on the aforementioned day.

The left column of maps displays the actual locations of congestion throughout the four time periods, while the right column depicts the predicted congestion locations. Congestion status is represented by three colors: blue indicates no congestion, orange indicates light congestion, and red denotes heavy congestion. Analyzing the ground truth congestion maps, we observe that congestion predominantly occurs along the east-west direction of the freeway, near the Louisiana coastline. This observation suggests that the primary evacuation routes are directed either westward towards Texas or eastward towards Mississippi. Temporally, congestion initiates during the morning period (6 AM to 12 PM), intensifies during the afternoon period (12 PM to 6 PM), and persists into the evening period. This distinctive pattern differs from daily recurring traffic congestion, typically concentrated within city roadways during peak commuting hours. In contrast, hurricane-induced congestion can manifest on freeways connecting cities and endure for extended hours.

Examining the prediction results shown in the right column maps, we find that the MLP model successfully identifies congested roadway segments during each time period, closely aligning with the actual congested locations. The predicted congested links can cover more than 85% of actual congested links across different time periods on the day before hurricane Ida made its landfall. However, it is worth noting that the predicted results occasionally show heavy congestion labels on road segments with either no congestion or only light congestion. This discrepancy suggests that the model may have overfitted the heavy congestion label compared to the other two labels, which corresponds to the recall results shown in Table 6.5.

In conclusion, our long-term congestion pattern prediction model, based on the MLP architecture, effectively captures and predicts the spatial distribution of hurricane-induced congestion during evacuation. It demonstrates a comprehensive understanding of congestion patterns across various time periods on the day before hurricane Ida’s landfall. Nonetheless,



some overfitting issues were observed, especially concerning heavy congestion predictions. As a result, further refinement and optimization of the model may be required to enhance its performance and accuracy in predicting congestion levels during such critical events.

### 6.5.2 Short-term traffic speed prediction

Table 6.5 presents a comprehensive comparison of prediction performance among different models under various prediction horizons. It is important to note that all models demonstrate high accuracy when dealing with recurring patterns and non-congested links. However, for the evaluation in Table 6.5, we focused exclusively on links experiencing heavy congestion throughout the 7-day hurricane period to assess their true performance in challenging scenarios.

Table 6.5: Performance of LSTM and other baseline models on long-term congestion prediction (average  $\pm$  standard deviation across 5 experimental repeats)

Model	Horizon (hour)	Metrics		
		RMSE (mi/h)	MAE (mi/h)	MAPE (%)
ARIMA	1	7.61 $\pm$ 0.31	6.94 $\pm$ 0.93	12.82 $\pm$ 1.86
	3	11.34 $\pm$ 0.83	13.93 $\pm$ 1.15	15.15 $\pm$ 4.07
	6	16.29 $\pm$ 1.27	17.73 $\pm$ 4.29	25.74 $\pm$ 6.95
Vanilla RNN	1	5.27 $\pm$ 0.66	3.75 $\pm$ 0.41	10.72 $\pm$ 2.54
	3	9.41 $\pm$ 0.27	10.83 $\pm$ 0.39	12.72 $\pm$ 3.21
	6	12.35 $\pm$ 1.71	14.61 $\pm$ 1.70	20.76 $\pm$ 5.03
LSTM	1	<b>2.72 <math>\pm</math> 0.06</b>	<b>2.12 <math>\pm</math> 0.16</b>	<b>7.72 <math>\pm</math> 2.73</b>
	3	<b>6.72 <math>\pm</math> 0.46</b>	<b>8.76 <math>\pm</math> 0.46</b>	<b>9.72 <math>\pm</math> 3.09</b>
	6	<b>8.76 <math>\pm</math> 0.80</b>	<b>8.03 <math>\pm</math> 0.50</b>	<b>13.76 <math>\pm</math> 4.80</b>

From the results presented in Table 6.5, it is evident that the LSTM model consistently outperforms the other two models across all three metrics. As the prediction horizon increases, all models experience a decrease in prediction accuracy, with the errors between predicted and true speed values becoming more pronounced. Notably, the LSTM model exhibits the remarkable ability to minimize prediction errors even when dealing with a long 6-hour prediction horizon. Interestingly, when comparing the performance of LSTM with Vanilla RNN and ARIMA, it becomes apparent that the LSTM model’s 6-hour prediction performance

surpasses the 1-hour and 3-hour predictions of the other two models. This finding suggests that the LSTM model effectively captures sequential dependencies between previous and future timesteps, enabling it to discern correlations between hurricane features and speed patterns more adeptly than the baseline models.

The results in Table 6.5 underscore the superiority of the LSTM model in tackling challenging congestion scenarios and extending its predictive capabilities to longer horizons. The LSTM’s ability to understand and leverage sequential dependencies enables it to excel in learning the intricate relationships between hurricane features and speed patterns, setting it apart from the Vanilla RNN and ARIMA models. These findings solidify the LSTM model’s potential as a valuable tool for traffic prediction tasks, especially during hurricane-induced congestion periods.

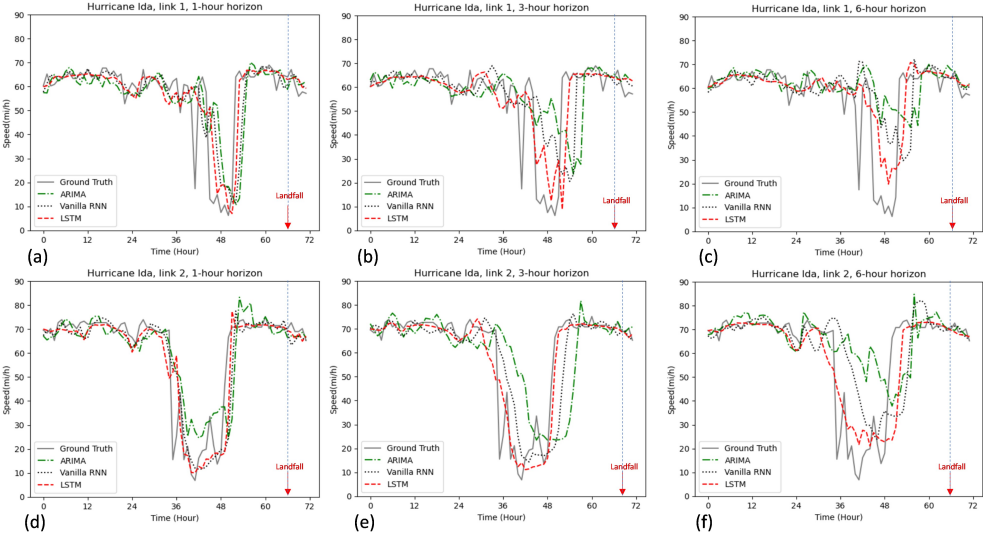


Figure 6.8: Comparison of prediction performance with LSTM and baseline models during hurricane Ida ((a), (b), (c): test link 1 with 1-, 3-, and 6-hour horizon; (d), (e), (f): test link 2 with 1-, 3-, and 6-hour horizons)

To intuitively present the performance of the short-term traffic speed prediction model, we selected two test links and plotted the continuous speed prediction over a multi-day range. As seen in Fig 6.6, link 1 is on northbound I-59 near the border with Mississippi, and link 2 is on westbound I-10 near the border with Texas. Both links are on major evacuation routes

of Louisiana during hurricane Ida. Fig 6.8 illustrates the prediction performance of LSTM and baseline models on two selected test links over a three-day range, from two days before hurricane Ida made its landfall to the day of landfall. The top three plots (Fig 6.8 (a), (b) and (c)) represent the results for test link 1, while the bottom three plots (Fig 6.8 (d), (e) and (f)) correspond to test link 2. Both tests were conducted using prediction horizons of 1, 3, and 6 hours. Several key findings emerged from the analysis:

- LSTM predictions exhibit reduced delay compared to ground truth when compared to baseline models. For a 1-hour prediction horizon, all three models demonstrate minimal delay in predicting congestion periods. As the prediction horizon increases, the LSTM model continues to accurately capture the start and end of congestion, while the baseline models show a substantial delay. Specifically, when the prediction horizon is extended to 6 hours, the ARIMA model can barely predict the onset of congestion, and the Vanilla RNN model's predicted onset time is 4 to 5 hours later than the actual congestion. A similar trend is observed when predicting the recovery of traffic speed from congestion. Both the ARIMA and Vanilla RNN models struggle to accurately predict the return of speed, with predicted recovery times being 2-3 hours late using a 3-hour prediction horizon and 5-6 hours late using a 6-hour prediction horizon.
- The LSTM model also demonstrates superior performance in predicting the lowest speed during periods of heavy congestion. For a 1-hour prediction horizon, both LSTM and Vanilla RNN models provide reasonably accurate predictions for the lowest speed. However, as the prediction horizon increases, the performance of the baseline models significantly deteriorates, while the LSTM model maintains its accuracy with only around a 5 to 10 mi/h error on the test links. Notably, when the horizon is set to 6 hours, the speed difference between the real and predicted lowest speed is approximately 30 mi/h for the Vanilla RNN model and 45 mi/h for the ARIMA model, in contrast to the LSTM's 10 mi/h difference. These findings suggest that the baseline models

are inadequate in providing valid predictions, not only in speed magnitude but also in the timing of congestion when the prediction horizon is relatively large. Additionally, the results demonstrate that the LSTM model effectively handles the spatial-temporal relationships required for short-term speed prediction, even when predicting speed values several hours into the future.

In conclusion, the LSTM model outperforms the baseline models in terms of reduced delay, accurate prediction of congestion periods, and the ability to predict lowest speeds during heavy congestion periods. The findings highlight the LSTM model’s capability to handle spatial-temporal dependencies in short-term speed prediction, making it a promising approach for traffic forecasting applications.

## 6.6 Conclusion

In this study, we discussed and defined the research question of network-level traffic speed prediction during hurricane evacuation with limited traffic data scenarios. A comprehensive model framework adopting the MLP and the LSTM is developed to learn the long-term congestion pattern and short-term speed pattern during hurricane evacuation. A case study using the Louisiana evacuation route network and archived speed data from 5 historical hurricanes demonstrated that the MLP long-term congestion state prediction achieved about 82% accuracy in predicting the congestion state of 6-hour period across the 7-day horizon. Additionally, the short-term speed prediction model achieved prediction MAPE from 7% to 13% for different horizons, ranging from 1 hour to 6 hours. Notably, both the MLP and LSTM outperformed other baseline models in terms of prediction accuracy. The case study results demonstrate that the proposed model framework presents a valuable contribution for efficient traffic management by offering a holistic approach to predict traffic conditions during hurricane events in a large-scale transportation network.

As for the future research directions, firstly, it is necessary to extend the current horizon for both MLP and LSTM and explore the upper bound of the prediction performance with extended horizons. Secondly, while the current model focuses only on traffic speed due to source data limitation, we should consider incorporating the output layer with additional traffic-related features if sufficient traffic data becomes available in the future.

# Chapter 7

## Leveraging data-centric AI for work zone traffic impact prediction

Work zone is one of the major causes of non-recurrent traffic congestion and road incidents. Despite the significance of its impact, studies on predicting the traffic impact of work zones remain scarce. In this study, we propose a data integration pipeline that enhances the utilization of work zone and traffic data from diversified platforms, and introduce a novel deep learning model to predict the traffic speed and incident likelihood during planned work zone events. The proposed model transforms traffic patterns into 2D space-time images for both model input and output and employs an attention-based multi-context convolutional encoder-decoder architecture to capture the spatial-temporal dependencies between work zone events and traffic variations. Trained and validated on four years of archived work zone traffic data from Maryland, USA, the model demonstrates superior performance over baseline models in predicting traffic speed, incident likelihood, and inferred traffic attributes such as queue length and congestion timings (i.e., start time and duration). Specifically, the proposed model outperforms the baseline models by reducing the prediction error of traffic speed by 5% to 34%, queue length by 11% to 29%, congestion timing by 6% to 17%, and increasing the accuracy of incident predictions by 5% to 7%. Consequently, this model offers substantial

promise for enhancing the planning and traffic management of work zones.

## 7.1 Introduction

The rising load on road infrastructures driven by population growth has resulted in an increased demand for road maintenance and reconstruction activities [75]. These work zone events often involve lane closures that lead to more traffic crashes and delays caused by reduced road capacity. Unlike the usual congestion seen during peak traffic hours, work zone activities typically create non-recurring, unexpected travel delays. According to the Federal Highway Administration [78], work zones account for nearly 24% of non-recurring traffic congestion. Additionally, work zone events significantly endanger the safety of both travelers and workers; for instance, in 2022, traffic accidents in work zones resulted in 891 fatalities [78].

To address safety and mobility requirements during highway maintenance and construction, and to align with the expectations of the travelers, it is important for traffic management and work zone planning agencies to have an accurate estimation of how work zone events will impact traffic. Modeling and predicting work zone impacts can enhance an agency's decision-making as well as its overall understanding of the factors affecting work zone decisions [76].

Research on predicting the impact of work zones on traffic is limited. Over the past few decades, related studies can be broadly divided into two categories: simulation or parametric-based approaches and non-parametric, data-driven approaches.

In the field of simulation-based studies, Ping and Zhu [252] estimated the changes in traffic capacity under various work zones using CORSIM. Chatterjee et al. [44] incorporated drivers' behavior into simulations and developed a work zone traffic flow estimation model in VISSIM. Wen [265] developed a work zone traffic simulation model for connected traffic conditions. These simulation-based models usually only consider a few work zone factors and network configurations, thus are mostly unable to predict traffic conditions under unseen

work zones with complex spatial-temporal patterns.

As the availability of data expands, facilitated by development in sensors and data collection techniques, the focus of research is increasingly turning toward data-driven methods, even though these data are not yet fully integrated. On the data-driven side, Adeli and Jiang [5] created a neuro-fuzzy model to estimate the traffic flow impacted by work zones. The results demonstrated the model’s superiority over empirical approaches. Karim and Adeli [128] proposed an adaptive neural network model to predict the traffic impact including capacity, queue length, and delay during work zones. Hou et al. [277] developed four machine learning-based work zone traffic prediction models: random forest, baseline predictor, regression tree, and neural network, evaluated on two selected roadway segments in St. Louis, MO, USA. Bae et al. [115] developed a multi-contextual machine learning method to model the traffic impact of urban highway work zones. By adopting machine learning-based approaches, these models handle more complex work zone conditions compared to simulation-based models. However, the performance is still constrained due to overly simplified model assumptions and structures. These models either provide only aggregated traffic indicator predictions or focus narrowly on specific aspects of the traffic impact caused by work zone events.

Reviewing existing research highlights two major limitations in predicting the impact of work zone traffic: 1) The quality and quantity of data sources are limited, as there is often no comprehensive pipeline for integrating, curating, and augmenting work zone traffic data for enhanced data-driven methods; 2) The data-driven methods currently used are overly simplified and not capable of handling the complex and dynamic traffic variations associated with work zones. Consequently, there is a pressing need for a model that can effectively capture the dependencies between spatial-temporal traffic patterns and work zone characteristics, providing a holistic perspective on both mobility and safety impacts.

The Work Zone Data Exchange (WZDx) Feed Registry, maintained by the U.S. Department of Transportation (USDOT), contains up-to-date metadata on work zone feeds that adhere to WZDx specifications [77]. Launched in 2019 by the Federal Highway Administration (FHWA)



and the Intelligent Transportation Systems Joint Program Office (ITS JPO), this initiative seeks to enhance road safety and mobility by standardizing work zone data and ensuring its broad accessibility in a consistent format [249]. In this study, we utilize WZDx datasets from the ITS DataHub combined with the University of Maryland CATT Laboratory’s Regional Integrated Transportation Information System (RITIS) data [249], providing insights into travel times and traffic speeds across Maryland’s transportation network. Additionally, we integrate the Maryland Department of Transportation’s (MDOT) Annual Average Daily Traffic statistics and loop detector data with incident data from RITIS and MDOT to create an enriched work zone dataset for predictive model training.

Besides the integration of multi-context datasets, the selection of data-driven models is crucial for estimating the traffic impact of work zones. Generally, traffic prediction models are categorized into short-term and long-term traffic forecasts [275] [12] [159] [116]. Both use a sequence-to-sequence or sequence-to-one approach, where a sequence of past traffic readings from the previous  $N$  timesteps is used to predict the traffic status for the following one or several timesteps, ranging from several minutes to multiple hours. These methods depend on the most recent traffic data to forecast future traffic conditions. However, these sequence-based models do not align with the objectives of our study. Our research aims to predict the traffic conditions on road segments with planned work zones well in advance before their implementation (e.g., days or weeks ahead), meaning no real-time traffic data at the time of making the prediction. To the best of the authors’ knowledge, none of the existing traffic prediction models are designed for such goals. Inspired by image-based sequence-to-sequence traffic prediction methods [181] [57] [206], which transform city-level space-time traffic states into 2D images for model inputs and outputs, this study proposes a novel image-to-image prediction method for work zone traffic forecasts. This method converts the historical spatial-temporal traffic patterns into multi-channel image inputs and conducts a joint representation with the planned work zone features to deliver a comprehensive traffic prediction for the entire duration of the work zones at once, which enables the estimation of

the traffic impact with high time resolution for the planning of work zones.

In summary, based on the curated dataset created by the data integration pipeline, we introduce an attention-based multi-context convolutional encoder-decoder neural network, named AMCNN-ED, to predict the impact of planned work zones, specifically focusing on mobility impacts such as traffic speed, queue length, congestion start time/duration, and safety impacts such as incident likelihood. The contributions of this study are outlined as follows:

- Developed a data curation pipeline that integrates work zone event data with traffic and roadway network datasets, creating an enhanced data source tailored for predicting the traffic impacts of work zone events.
- Introduced an image-based modeling approach to estimate traffic impact caused by work zones by converting historical space-time traffic patterns into 2D images as model inputs. Based on that, we developed a novel attention-enhanced multi-context convolutional encoder-decoder neural network structure to capture the spatial-temporal dependencies between work zone characteristics and dynamic traffic patterns, enabling in-advance prediction of traffic impact (i.e., speed, queue length, congestion start time/duration, and incident likelihood) for planned work zones well ahead of time.
- Conducted a comprehensive evaluation of the proposed model using a real-world dataset from Maryland’s transportation network, benchmarking it against baseline models to demonstrate its superior performance.

## 7.2 Data Integration and Reconstruction

### 7.2.1 Multi-Context Work Zone Data

To construct a comprehensive work zone dataset, this study follows the process outlined in Fig. 7.1, which details the integration of datasets from various sources. WZDx provides dynamic and detailed work zone information, including precise locations, timings, lane counts, geometries, and potential vehicular impacts. Traffic data from RITIS adds link-level metrics such as travel time and traffic speed, along with historical and reference speeds, enriching the dataset further. By combining this data with MDOT’s volume data and supplementing it with incident information, the dataset achieves a high level of granularity, encompassing individual work zones and their broader impact on the transportation network.

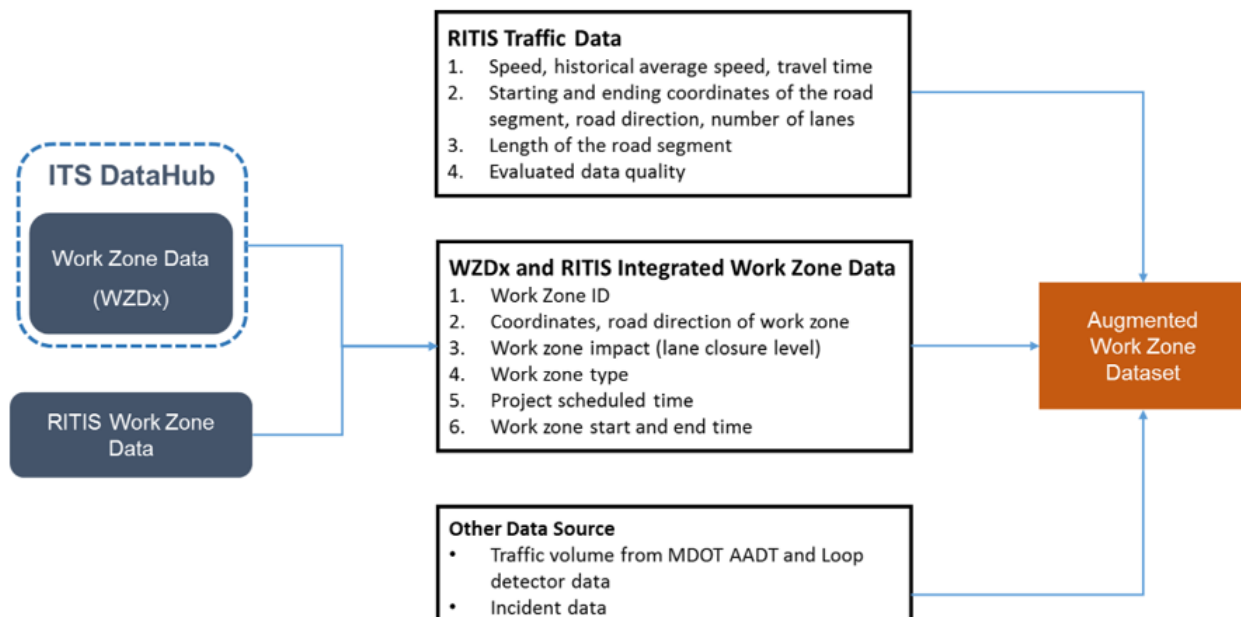


Figure 7.1: Multi-context data integration pipeline

By incorporating these diverse data streams, valuable operational metrics are acquired at a granular level, encompassing individual work zones, their immediate surroundings, the impacted corridor, and the broader regional road network. Furthermore, the robust

data capabilities will also facilitate more in-depth categorization based on different types of work zones and specific geographical regions. This enhanced categorization will provide an understanding of the diverse impacts and dynamics across various work zone scenarios and geographic contexts.

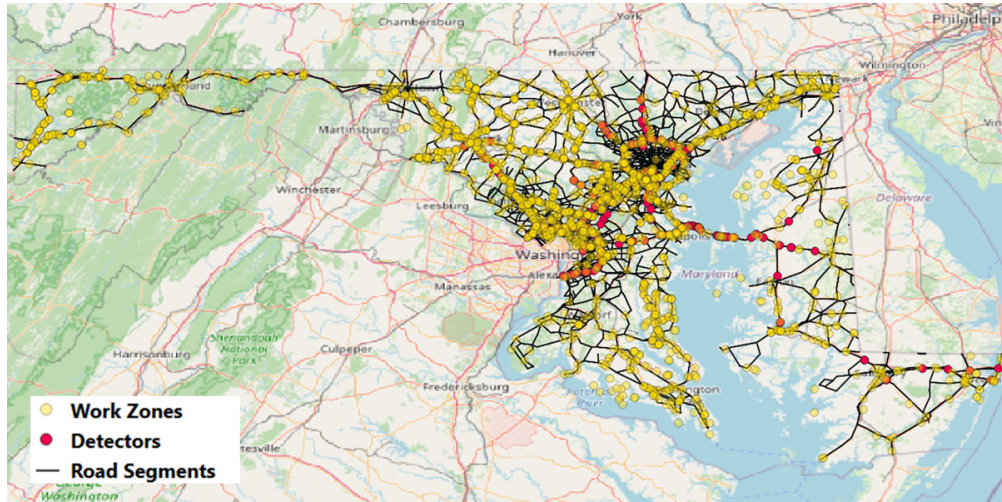


Figure 7.2: Spatial distribution of work zone, detectors, and road segments in Maryland transportation network

## 7.2.2 Data Integration and Space-Time Traffic Image Generation

The integration of these datasets employs a sophisticated spatial-temporal matching process. As illustrated in Fig 7.2, this map highlights the geospatial alignment of work zones, loop detectors, and road segments throughout the Maryland transportation network. Initially, matching is conducted using precise GPS coordinates to ensure each work zone is accurately paired with its corresponding road segment. Subsequently, traffic data of road segments and loop detectors relevant to the operational hours of each work zone are extracted. This dual-layered matching strategy—first spatial, then temporal—ensures a seamless amalgamation of location and time-specific traffic patterns.

Specifically, after cleaning and filtering, a total of 3646 work zones were identified from 2016 to 2019, excluding 2020 to 2022 due to the biased traffic patterns during the COVID-19

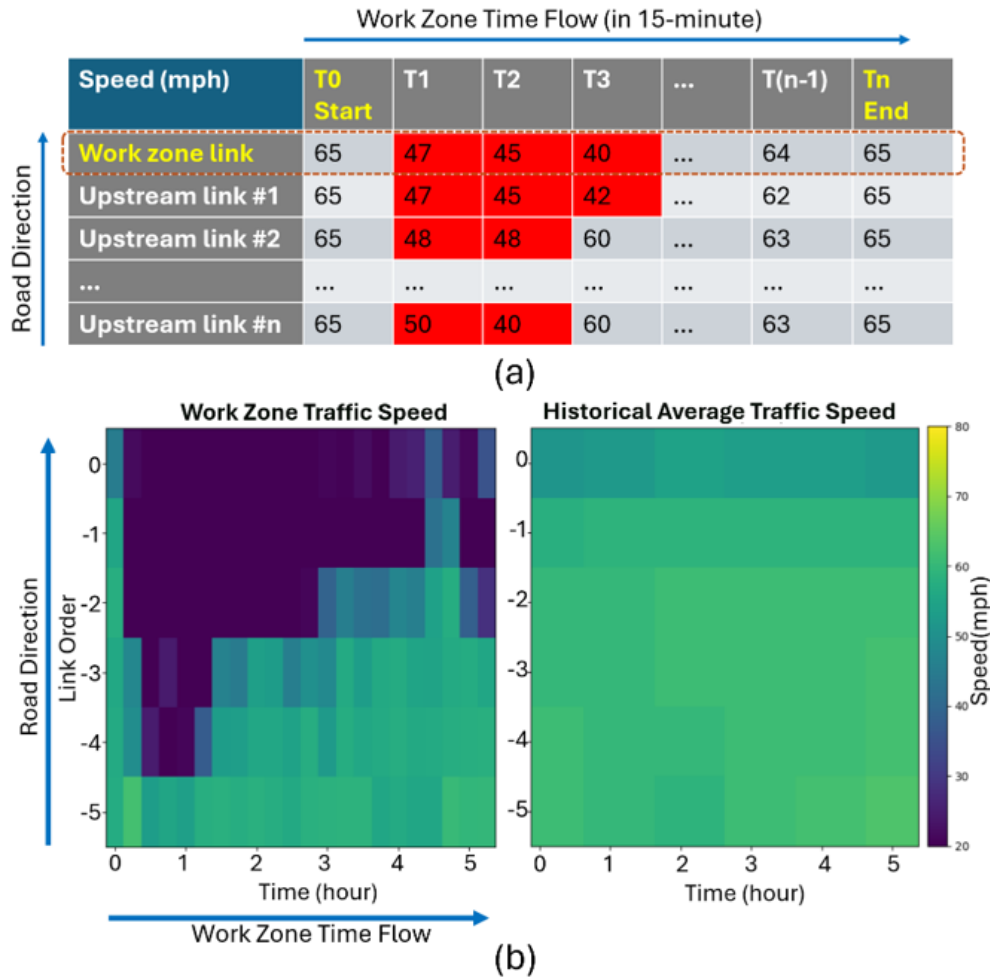


Figure 7.3: Work zone space-time traffic image generation: (a) 2D space-time traffic matrix; (b) Two examples of converted 2D space-time traffic images (work zone traffic speed and historical traffic speed)

pandemic. The study focuses on temporary work zones with durations of less than 24 hours. To capture pertinent data for the temporal work zone study and enable the development of an AI model, our approach compiles data on a case-by-case basis for each work zone. For every individual case, we systematically collect traffic data encompassing the complete duration of the work zone event. In terms of spatial information, we include data for road segments extending 5 miles upstream relative to each work zone. Based on this spatial matching result, a feature of “distance to work zone” is calculated for each road segment; similarly, “time to work zone start” and “time after work zone end” are calculated for each time step. To ensure a high level of data fidelity, we maintain a time resolution of 15 minutes throughout the dataset.

As a result, for each work zone case, as shown in Fig 7.3 (a), a 2D space-time matrix containing spatial-temporal information is organized, with the highlighted area indicating affected traffic. This matrix can be used to further represent other traffic features such as speed, historical average speed, and historical average volume, as well as geospatial features like link length and distance to the work zone link, all updated in 15-minute intervals. These 2D space-time matrices are further converted to 2D heatmap images with different colors indicating different levels of values for particular traffic features, as seen in Fig 7.3 (b). This systematic organization not only captures the real-time dynamics of work zones but also furnishes a standardized dataset format from which AI models can learn the complex spatial-temporal dependencies of traffic flow in relation to work zone activities, enhancing predictive capabilities.

## 7.3 Methodology

### 7.3.1 Problem Definition

The methodology proposed by this study tackles the problem of spatial-temporal traffic speed and incident likelihood prediction on road segments of planned work zone events well in advance before their implementation (e.g., days or weeks ahead). The definition of this predictive problem is presented as follows: For a planned work zone event scheduled to start at  $T_0$  and end at  $T_n$  at location  $L$  of a roadway, we define all the link segments on the same roadway within 5 miles upstream of  $L$  as target links. Assume that for these target links, the historical average traffic sequences (e.g., traffic speed, volume) during the same time of day and day of week corresponding to the planned work zone schedules are known. Additionally, the geospatial correlations between the links (e.g., link length, link order) and the characteristics of the planned work zone event (e.g., number of closed lanes, number of total lanes, etc.) are also known. The model aims to predict two key outcomes: 1) A sequence of traffic speeds on all the target links throughout the duration of the planned work zone; and 2) The likelihood of an incident occurring on the target links during the work zone period. The spatial-temporal traffic speed output can further be used to infer other traffic impact attributes such as maximum queue length, congestion start time, and congestion duration.

### 7.3.2 Model Structure

#### Model Overview

As shown in Fig 7.4, the model input encompasses two components: 1) A set of historical traffic patterns and geospatial sequences, which have been converted into a multi-channel 2D space-time image, and 2) A tabular feature vector of work zone characteristics. The 2D space-time image consists of multiple channels, each representing the historical traffic pattern and geospatial correlations of the link segments within the 5-mile range upstream of the work

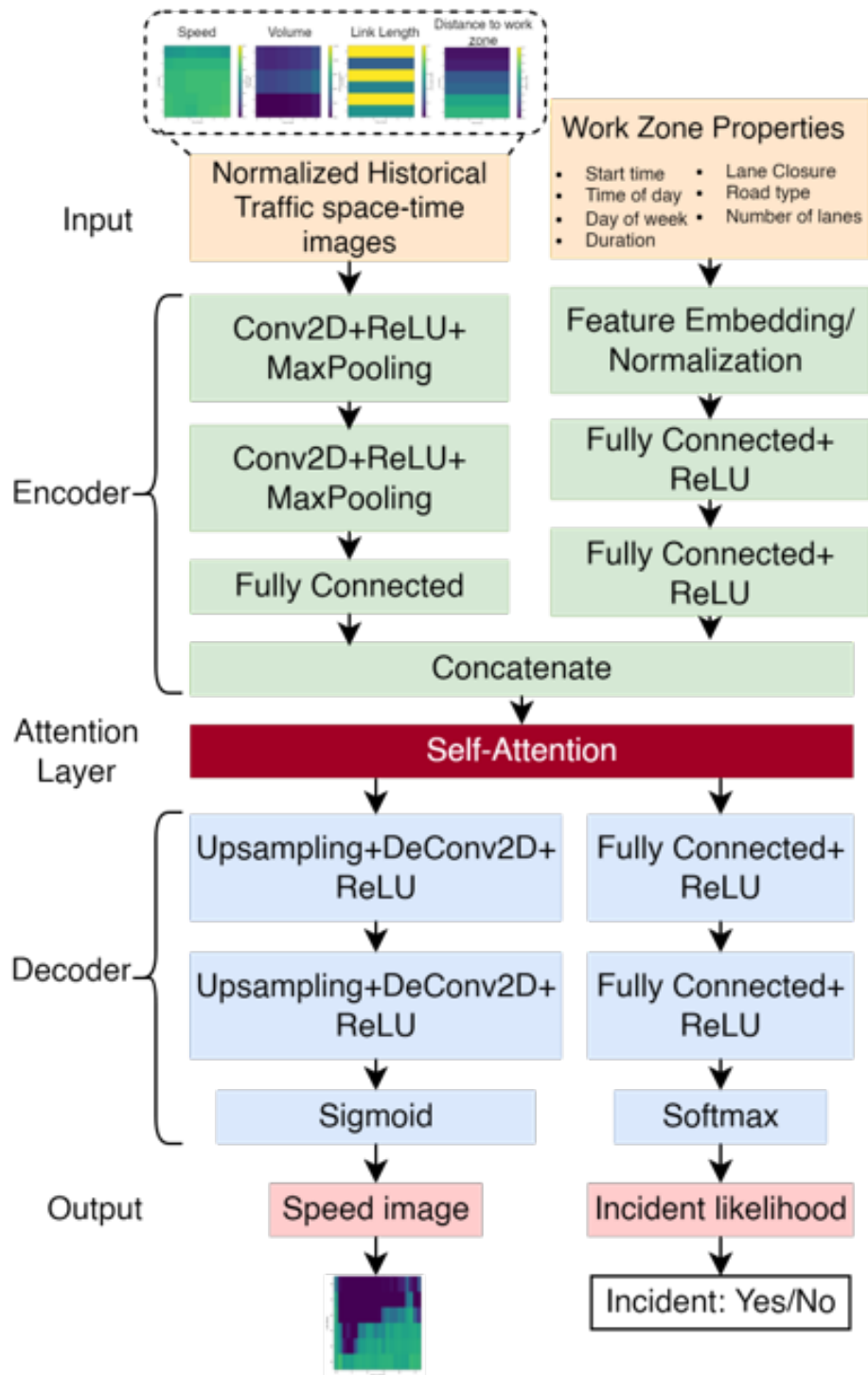


Figure 7.4: Model structure of AMCNN-ED



zone. The input image,  $X_{\text{image}}$ , can be defined as:

$$X_{\text{image}} \in \mathbb{R}^{h \times w \times c} = \{I_1^{h \times w}, I_2^{h \times w}, \dots, I_c^{h \times w}\} \quad (7.1)$$

where  $h$  refers to the height of the image (the number of links within the work zone 5-mile range),  $w$  refers to the width of the image (the number of timesteps of the work zone event), and  $c$  is the number of channels of the input image, representing features related to historical traffic patterns and geospatial relationships. In this study, we selected historical average speed, historical average volume, link length, and distance to the work zone location as the four channels of the input image.

The second input component is the feature vector of planned work zone characteristics, denoted as:

$$X_{\text{wz}} \in \mathbb{R}^{n \times 1} = \{x_1, x_2, \dots, x_n\} \quad (7.2)$$

where  $n$  denotes the number of features of the work zone. In this study, we consider the following features: start time of day, day of the week, work zone duration, number of lanes closed, number of total lanes, road type, and on-ramp/off-ramp connection.

The output of the model is defined as:

$$Y = \{Y_{\text{speed}}, Y_{\text{inci}}\} \quad (7.3)$$

This includes a predicted 2D space-time traffic speed image  $Y_{\text{speed}} \in \mathbb{R}^{h \times w \times 1}$ , which indicates the speed of target links within 5 miles upstream of the anticipated work zone at 15-minute intervals for the work zone duration, and a likelihood indicating the probability of incident occurrence during the projected work zone event. The goal of the work zone traffic impact prediction is to learn a mapping function  $f : X \rightarrow Y$  that can predict the traffic speed for each timestep on each upstream link and the likelihood of incident occurrence during the

work zone event.

The AMCNN-ED model constructs 3 modules: encoder layers, attention layer, and decoder layers. The multi-context encoder extracts the spatiotemporal features from historical space-time traffic data and static work zone features from planned work zone tabular data. The extracted feature maps are combined and passed to the attention layer to weigh the importance of each part in the concatenated feature representation. Then the attention-enhanced feature vector is sent to decoder layers with multiple transposed CNN layers and split in the output layer to generate both the 2D speed image and incident likelihood.

### **Encoder Layers**

The encoder consists of two parallel modules designed to create a joint representation of historical traffic information, geospatial features, and planned work zone characteristics. As shown in Fig 7.4, the image encoder module employs two convolutional neural network (CNN) layers to extract spatial-temporal dependencies in the historical traffic patterns of upstream links during the work zone period. Each convolutional layer comprises a 2-dimensional convolution layer (Conv2D), a ReLU activation layer, and a max pooling layer, which collectively extract spatial-temporal features from the preceding layer. At the end of the two CNN layers, a flattened layer converts the feature map into a 1D vector representation. Additionally, a tabular feature extraction module extracts features from work zone-related attributes and converts them into a 1D feature vector, which can then be concatenated with the feature vector extracted from the CNN layer.

### **Attention Layer**

As presented in Fig 7.4, the proposed network uses the self-attention mechanism to weigh the importance of different parts of the feature representation from the encoder layer. The self-attention mechanism is a deep learning technique originally designed for natural language processing (NLP) tasks to improve the modeling of relationships in sequential data [2], and

further implemented in other areas such as helping the model to learn which part of the feature representation is more informative for succeeding model components [232, 285].

When image and tabular data features are concatenated, they form a combined feature space. This space includes both the spatial information from the images and structured information from the tabular data. However, not all features contribute equally to the task at hand. An attention mechanism is employed here to dynamically learn to focus more on those features that are more relevant, effectively learning a task-specific weighting of features. By applying attention to the concatenated features, the model can highlight aspects of the data that are more informative for the specific prediction or reconstruction task. This selective focus can improve accuracy and robustness by reducing the impact of less relevant or noisy data.

To compute the decoder input, first, features extracted from both the image and the tabular data are combined into a single feature vector. This combined feature vector is then transformed into three different sets of vectors [232]: queries ( $Q$ ), keys ( $K$ ), and values ( $V$ ). These transformations are achieved through multiplication by three distinct sets of weights. The model computes scores by comparing all the queries with all the keys. These scores determine how much attention or importance should be given to each value vector. Each value vector is then multiplied by its corresponding attention score, effectively emphasizing more important features and diminishing less important ones. The resulting weighted sum forms a new, attention-enhanced feature vector that is used as the input for the decoder. The process of implementing the self-attention mechanism on the encoded input feature can be expressed by the following equations:

$$\left\{ \begin{array}{l} [Q, K, V] = [W_Q, W_K, W_V] \cdot x \\ A = \text{softmax} \left( \frac{Q \cdot K^T}{\sqrt{d_k}} \right) \\ z = AV \end{array} \right. \quad (7.4)$$

where  $x$  is the concatenated input feature vector from the encoder layer,  $W_Q, W_K$ , and  $W_V$  are weight matrices,  $\sqrt{d_k}$  is a scaling factor, and  $z$  is the output feature vector after applying self-attention.

## Decoder Layers

The decoder layer of this network consists of two transposed convolution layers for image reconstruction and a set of fully connected layers for incident likelihood prediction. The transposed convolution layers, denoted as DeCNN, are used to reconstruct the encoded feature vector to produce a 2D image for speed prediction.

The decoder receives an attention-enhanced feature vector from the attention layer. Then, the first DeCNN layer takes the flattened feature vector from the previous layer and reshapes it back into a multi-dimensional tensor. It then applies transposed convolution operations to start upsampling the features back to the spatial dimensions needed for image reconstruction. Following the initial upsampling, the second DeCNN layer further increases the spatial dimensions of the feature map, continuing to add detail and structure. It reduces the number of channels, aiming to reconstruct the spatial structure of the original input image. After each transposed convolution, an activation function such as ReLU is applied to introduce non-linearity, helping to model complex patterns in the data.

Following two consecutive DeCNN layers, the image output path employs a sigmoid activation to normalize the image pixels for the one-channel speed graph. On the classification output side, the attention-enhanced feature vector is sent to a set of fully connected layers with a softmax activation function at the end to output a probability between 0 and 1, indicating the likelihood of the input belonging to one of two incident labels.

### 7.3.3 Loss Function

Given that the model adopts a multi-task learning structure and outputs two types of outputs, i.e., 2D space-time traffic speed image and incident likelihood, we employ distinct loss functions for each target output and combine them to represent the model’s overall loss.

For traffic speed prediction, we implement the widely used Huber loss function to mitigate the impact of outliers in speed predictions [158]. The definition of Huber loss is provided in Equation 7.5, where  $y$  and  $\hat{y}$  refer to observed and predicted speeds, respectively, and  $\delta$  is a hyperparameter that requires tuning:

$$L_{\delta}(y, \hat{y}) = \begin{cases} \frac{1}{2}(y - \hat{y})^2 & \text{for } |y - \hat{y}| \leq \delta \\ \delta|y - \hat{y}| - \frac{1}{2}\delta^2 & \text{otherwise} \end{cases} \quad (7.5)$$

For incident prediction, we employ cross-entropy loss, commonly used in classification problems [115], denoted as  $L_{ce}$ . This loss function measures the performance of the classification output, which is a probability value between 0 and 1. The total loss can be expressed as the weighted sum of the losses from the two tasks:

$$L = w_1 \cdot L_{\delta} + w_2 \cdot L_{ce} \quad (7.6)$$

where weights  $w_1$  and  $w_2$  are hyperparameters to be tuned during model training.

## 7.4 Experimentation

### 7.4.1 Performance Metrics

The experimentation adopts three widely applied evaluation metrics to quantify the performance of speed prediction of each model [2, 218, 219]. They are Mean Square Error (MSE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). The

performance metrics are presented in (5), where  $\hat{y}_i$  represents the predicted speed made by the model, and  $y_i$  represents the corresponding ground-truth value. The evaluation metrics are defined as:

$$\left\{ \begin{array}{l} RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2} \\ MAE = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| \\ MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|\hat{y}_i - y_i|}{y_i} \end{array} \right. \quad (7.7)$$

For incident prediction, we adopted three commonly used classification prediction metrics: recall, precision, and F1 score to assess each model’s performance [157]. Recall measures the proportion of actual positives correctly identified by the model, highlighting its sensitivity. Precision assesses the accuracy of the positive predictions made by the model, indicating the proportion of true positives among all positive predictions. F1 score is the harmonic mean of precision and recall, providing a single metric that balances both precision and recall to measure a model’s accuracy more comprehensively.

For work zone samples exhibiting congestion patterns, we introduced three congestion-specific metrics to evaluate the prediction performance: the start time, duration, and maximum queue length of the congestion, where the duration and queue length are the width and depth of the congestion area on the space-time speed image. For each 2D space-time image corresponding to a work zone, to minimize the interference of random data noise, we only consider congestion that lasts over one hour and extends across multiple consecutive link segments as valid. To identify valid congestion areas in the space-time images, we employed Otsu’s method, an automated process widely used in image segmentation. Otsu’s thresholding algorithm is particularly effective for automatically performing clustering-based image thresholding [230]. The method operates by calculating the histogram of pixel

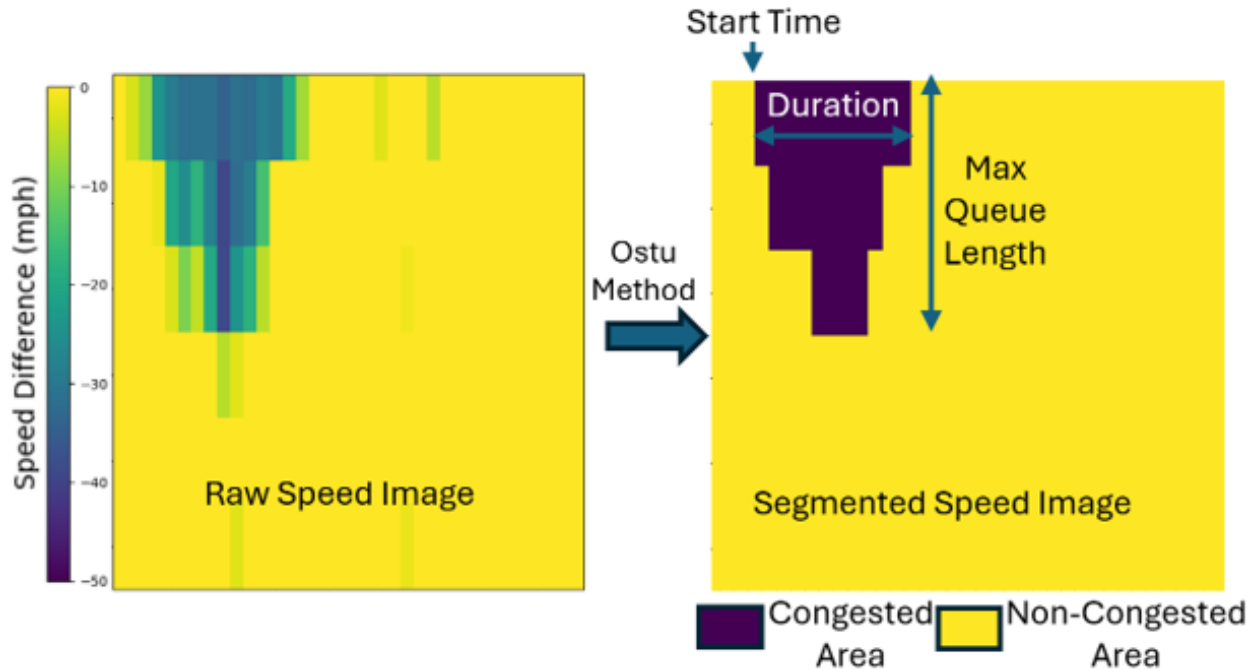


Figure 7.5: Work zone space-time traffic image processing

intensities and systematically testing all possible thresholds to determine which maximizes the between-class variance (i.e., the variance between the pixel intensities above and below the threshold) [274].

### 7.4.2 Baseline Models

The results of our model are compared against the following models:

- **ARIMA**: Auto-regressive integrated moving average.
- **GRU**: Gated recurrent unit network.
- **LSTM**: Long-short-term memory network.
- **Conv-LSTM**: Convolutional long-short-term memory network.

The first four models—ARIMA [16], GRU [220], LSTM [220], and Conv-LSTM [181]—all make speed predictions in an autoregressive form. They require a short initial sequence as

input to predict the very first timestep during the work zone. They then gradually append the newly predicted speed values to the input sequence and use the extended sequence to predict the next timestep until the entire duration of the work zone is predicted. The MCNN-ED model uses the same encoder-decoder structure as the AMCNN-ED proposed by this study, with the only difference being that MCNN-ED does not incorporate a self-attention layer to enhance the feature representation.

It should be noted that there aren't any existing models that can be applied directly for the problem defined in this study, therefore the baseline models listed here are highly customized to fit the work zone prediction scenario in this study. The literature cited here only provided high-level concepts instead of complete model structures.

All neural network models were implemented using Pytorch 2.0. Each model was trained on an RTX A5000 GPU, providing ample GPU memory to facilitate the learning process. Additionally, the Adam optimizer was employed. The models were run for 200 epochs, with early stopping implemented to prevent overfitting. Training was halted if the validation loss deteriorated for a specified number of epochs, even if the training loss continued to decrease.

## 7.5 Results and Performance Evaluation

### 7.5.1 System-Level Performance Analysis

Tables 7.1, 7.2, and 7.3 present the prediction results of the proposed model and baseline models on the testing dataset for the 547 work zone events. Table 7.1 displays the results for all test work zones, while Table 7.2 focuses on results in congested areas of impacted work zones. The results in Table 7.1 demonstrate that the neural network-based models all outperform the ARIMA model. This is because ARIMA relies solely on previous timesteps' traffic data and fails to account for changes in traffic caused by work zone activities. Additionally, the results indicate that RNN-based autoregressive models do not perform as well as encoder-decoder



structures, due to their inability to capture the comprehensive spatial-temporal dependencies between work zone properties and traffic patterns. Among the two encoder-decoder models, AMCNN-ED outperforms CNN-ED. This superior performance can be attributed to the self-attention layer in AMCNN-ED, which enhances the model’s ability to discern the relative importance of different sectors in the joint feature representations produced by the encoder layers.

Table 7.1: Speed Prediction Results

<b>Model</b>		<b>Performance metrics</b>		
		<i><b>MAE</b></i>	<i><b>RMSE</b></i>	<i><b>MAPE</b></i>
Auto- Regression	ARIMA	10.83	12.11	17.72
	GRU	8.14	8.63	14.65
	LSTM	7.99	8.61	14.53
	Conv-LSTM	7.59	8.46	13.83
Encoder- Decoder	MCNN-ED	7.36	8.30	13.77
	<b>AMCNN-ED</b>	<b>7.10</b>	8.06	13.16

From the perspective of in-advance traffic management and long-term work zone planning, accurately forecasting the road segments impacted by work zone activities is of paramount importance. Therefore, we selected 50 work zone samples that experienced congestion during the work zone duration from a total of 547 test work zones to compare the performance between our model and baseline models. The ARIMA model, unable to predict traffic congestion caused by work zone activities, was excluded from the analysis in Table 7.2.

As shown in Table 7.2, compared to RNN-based autoregressive models, the two encoder-decoder approaches demonstrate substantial improvements. This suggests that multi-context convolutional feature extraction is more effective at capturing the spatial-temporal correlations

Table 7.2: Congested Area Prediction Results

<b>Model</b>		<b>RMSE</b>		
		<i><b>Congestion Start Time</b></i>	<i><b>Congestion Duration</b></i>	<i><b>Max Queue Length</b></i>
Auto-Regression	GRU	2.32	2.97	1.53
	LSTM	2.24	2.99	1.45
	ConvLSTM	2.18	2.88	1.36
Encoder-Decoder	MCNN-ED	2.03	2.65	1.22
	<b>AMCNN-ED</b>	<b>1.92</b>	2.53	1.09

across multiple adjacent locations over extended periods. This capability is particularly crucial for predicting non-recurrent congestion patterns during work zone events. Furthermore, the AMCNN-ED structure outperforms the CNN-ED structure, primarily due to its self-attention mechanism, which enables the model to identify key elements in the feature vectors from both the static work zone features and the historical spatial-temporal traffic patterns, thus more accurately predicting the occurrence of traffic congestion.

Table 7.3 presents the prediction results for collision incidents during work zone events. We excluded ARIMA from the model list since it is designed solely for time-series prediction and cannot provide classification outputs. The results show that the AMCNN-ED model outperforms the baseline model across all three performance metrics. This suggests that the AMCNN-ED model is more effective at predicting potential collision incidents compared to autoregressive models and non-attention-based encoder-decoder models, while also minimizing false alarms in work zones. It should be noted that the prediction accuracy of all the listed models remains below 0.7. This limitation is largely due to the stochastic nature of incidents and the current limitations of available data. According to various studies [157, 165, 231], the

occurrence of collisions is influenced by numerous factors, including traffic, road closures, and external conditions such as weather, driver behavior, and vehicle conditions. Therefore, it is challenging to achieve precise forecasts for incident occurrences based solely on historical traffic data and projected work zone properties. However, the results demonstrated by this model still show promising potential to assist in the prevention of potential crashes during the planning of work zone activities.

### 7.5.2 Event-Level Spatial-Temporal Performance Analysis

In this section, we focus on the model performance of selected examples from the test dataset to illustrate the prediction performance of the proposed AMCNN-ED model and compare it with the best-performing baseline model at the event level, as shown in Fig 7.6. Each plot in Fig 7.6 represents a 2D space-time speed difference graph for the duration of each work zone. In Fig 7.6, the ground truth speed graph is displayed in the left column, the prediction results from AMCNN-ED are in the middle column, and the results from the best-performing autoregressive model (Conv-LSTM) are in the right column. The dark blue areas indicate significant speed drops compared to the historical average speed at the same time of day, signaling severe congestion, while the yellow areas indicate speeds similar to the historical average. As shown in Fig 7.6, the four work zones caused one or more instances of congestion during the work zone period, extending to multiple link segments upstream. The AMCNN-ED model more accurately captures the timing and spatial extent of the congestion compared to the autoregressive model. In contrast, the Conv-LSTM model tends to underpredict or overpredict the congestion area.

A key insight from this comparison is that autoregressive models may incorrectly interpret traffic flow’s temporal variations. This occurs because they predict each timestep based solely on previous timesteps, ignoring shockwave propagation. In contrast, the AMCNN-ED model incorporates global information, both temporally and spatially, for the work zone event.

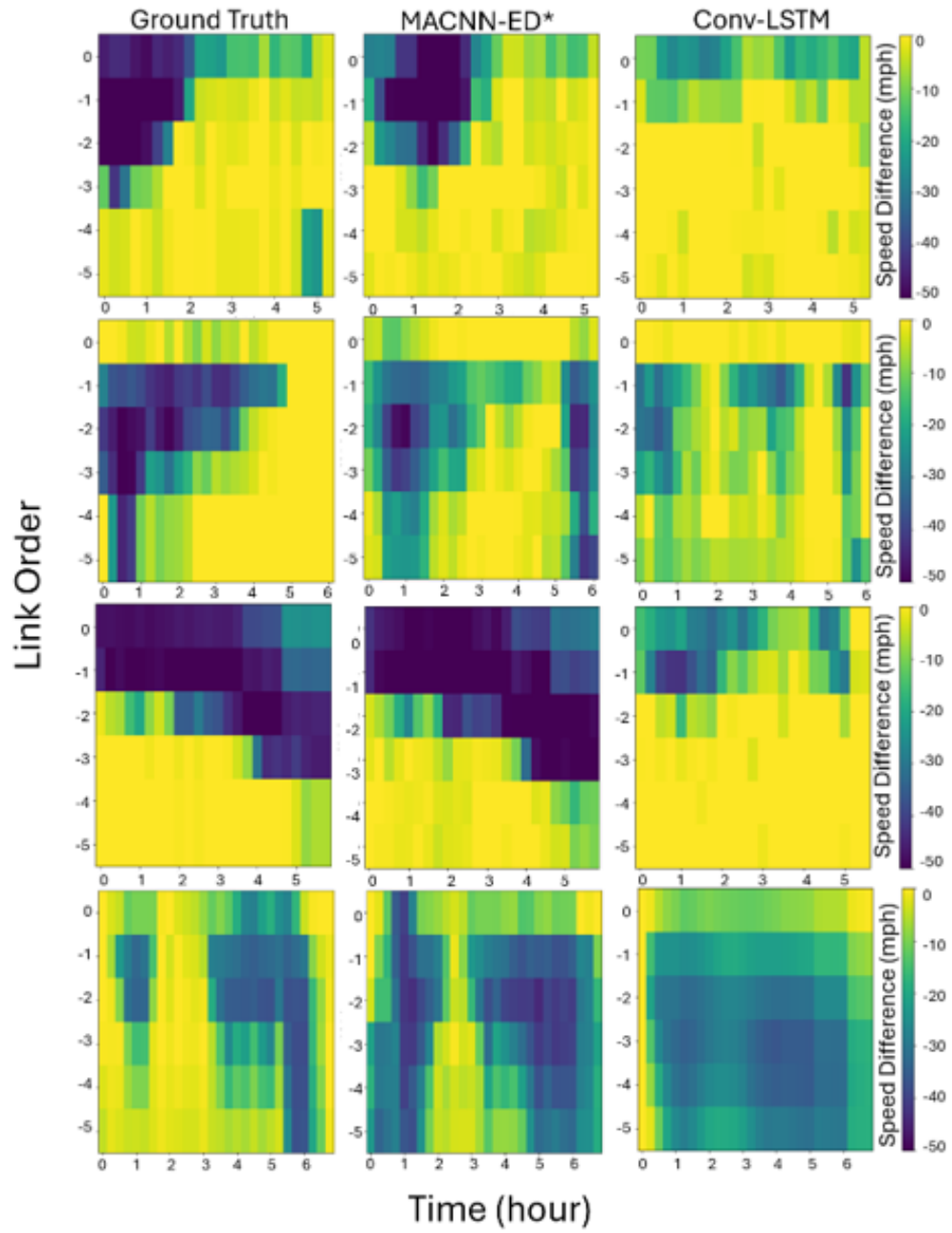


Figure 7.6: Examples of event-level speed prediction performance

Table 7.3: Incident Prediction Results

<b>Model</b>		<b>Performance metrics</b>		
		<i>Recall</i>	<i>Precision</i>	<i>F1 Score</i>
Auto-Regression	GRU	0.54	0.62	0.58
	LSTM	0.53	0.61	0.57
	ConvLSTM	0.54	0.62	0.58
Encoder-Decoder	MCNN-ED	0.56	0.62	0.58
	<b>AMCNN-ED</b>	<b>0.58</b>	0.65	0.61

This underscores the benefits of using an attention-based encoder-decoder structure over an autoregressive structure for predicting traffic patterns during planned work zone activities.

## 7.6 Conclusion

This study presents a data curation pipeline for data-centric work zone traffic prediction problems and proposes an attention-based multi-context encoder-decoder convolutional model to predict the traffic impact of planned work zone events. Our method consists of two main steps. First, we integrated archived data from multiple data platforms to construct a curated work zone traffic dataset that encompasses the essential factors influencing traffic changes and work zone characteristics. Next, we developed a convolutional encoder-decoder model to create a joint representation of multi-context spatial-temporal input features and implemented a self-attention mechanism to highlight key sectors within the encoded features. These features are then reconstructed through the transposed convolutional decoder layers to generate predictions for traffic speed and incident likelihood during the work zone events.

# Chapter 8

## Conclusions and Future Work

### 8.1 Conclusions

In the concluding remarks of this dissertation, we reflect on the advancements in using mobility agent and network modeling to create a comprehensive digital representation of transportation systems. This research explores both the synthetic and practical applications of digital twin frameworks for transportation systems, facilitating improved decision-making under dynamic mobility and traffic conditions. The key contributions of this dissertation are outlined as follows:

- We successfully developed a digital twin framework that integrates real-world mobility data with advanced simulations for decision support in transportation systems. By leveraging multi-context data from the physical transportation system and combining it with AI-driven mobility agents and network models, this framework addresses the limitations of traditional approaches, particularly in adaptability, transferability, and predictability. The digital twin framework offers both online and offline decision-making capabilities, providing actionable insights for real-time traffic management and long-term transportation planning.

- We introduced an agent-based modeling approach designed to integrate demand-side and supply-side variations for more adaptable transportation system models. The framework is tailored to accommodate new mobility trends, changing travel behaviors, and emerging vehicle usage patterns. By enhancing adaptability, this agent-based approach ensures that travel demand and network simulation models remain relevant and responsive to the evolving transportation innovations.
- We explored a fully data-driven AI agent modeling approach for synthesizing human mobility patterns. This approach automates the traditionally resource-intensive process of travel demand modeling, providing more efficient and transferable models. The AI agent models were evaluated within an agent-based network traffic simulation environment, demonstrating high fidelity and efficiency. This contribution significantly advances the automation of mobility pattern generation, making it applicable to regions with limited mobility data.
- Lastly, we successfully developed multiple AI network modeling approaches to improve the predictability of traffic states, especially during non-recurrent events like natural disasters or special road conditions. By incorporating multi-context data and leveraging deep learning models, this research enhances the accuracy of spatial-temporal traffic predictions. These models are critical for both regional transportation planning and real-time traffic management, providing decision-makers with better tools to manage and predict traffic flows in unpredictable scenarios.

In conclusion, this dissertation contributes to transportation systems research by integrating intelligent agents, advanced network models, and a digital twin framework, offering robust solutions for adaptability, transferability, and traffic predictability. These advancements support smarter decision-making processes, ensuring future transportation systems are equipped to handle the complexities of modern mobility and dynamic traffic conditions.

## 8.2 Future Work

Looking ahead, there are still several promising paths for advancing research in mobility agent and network modeling.

- Firstly, data from multiple contexts should be integrated into the synthesis of mobility patterns. Currently, mobility modeling heavily depends on travel survey data, which is often limited or outdated in many regions. To address this, the capability to extract meaningful insights from alternative data sources, such as social media activity and person-based GPS log data (e.g., Veraset), should be prioritized. Using more advanced AI techniques, such as large language models (LLMs), to interpret these diverse data streams could significantly enhance the richness, functionality, and robustness of both mobility agents and network models. With these multisource data inputs, models can capture a broader spectrum of human mobility patterns and adapt to shifts in behavior and context that are often missed by traditional data sources alone.
- Additionally, the dynamic integration of AI agents within network modeling deserves further attention. Advanced traffic prediction models, which go beyond traditional simulation processes in agent-based multimodal transportation, could help account for non-recurrent roadway conditions, such as accidents, weather changes, and events, that impact travel patterns. By incorporating predictive elements into the core framework of transportation system models, we can improve the overall accuracy and reliability of predictions across various real-world scenarios, offering more resilient and adaptable solutions for transportation management.
- Lastly, bridging the gap between academic research and practical implementation in transportation modeling remains a critical focus. By collecting more comprehensive and naturalistic travel behavior and traffic data, we can create models that more closely represent real-world interactions between agents and networks. This effort would



also involve closer collaboration with practitioners to ensure that models are not only theoretically sound but also practically viable and actionable in real-world applications, ultimately facilitating the deployment of innovative mobility solutions.

# Bibliography

- [1] Josm. <https://josm.openstreetmap.de>, 2020.
- [2] M. Abdel-Aty A. Abdelraouf and J. Yuan. Utilizing attention-based multi-encoder-decoder neural networks for freeway traffic speed prediction. *IEEE Transactions on Intelligent Transportation Systems*, 23(8):11960–11969, 2021.
- [3] A. Adebisi, Y. Guo, B. Schroeder, J. Ma, B. Cesme, A. Bibeka, and A. Morgan. Highway capacity manual capacity adjustment factor development for connected and automated traffic at signalized intersections. *Journal of Transportation Engineering, Part A: Systems*, 148(3):04021121, 2022.
- [4] A. Adebisi, Y. Liu, B. Schroeder, J. Ma, B. Cesme, A. Jia, and A. Morgan. Developing highway capacity manual capacity adjustment factors for connected and automated traffic on freeway segments. *Transportation Research Record*, 2674(10):401–415, 2020.
- [5] H. Adeli and X. Jiang. Neuro-fuzzy logic model for freeway work zone capacity estimation. *Journal of Transportation Engineering*, 129(5):484–493, 2003.
- [6] Federal Highway Administration. Integrated modeling for road condition prediction. U.S. Department of Transportation.
- [7] Federal Highway Administration. Integrated modeling for road condition prediction - phase 4 final evaluation report – louisiana, 2022.

- [8] T. Afrin and N. Yodo. A survey of road traffic congestion measures towards a sustainable and resilient transportation system. *Sustainability*, 12(11):4660, 2020.
- [9] J. Ahn, E. Ko, and E. Y. Kim. Highway traffic flow prediction using support vector regression and bayesian classifier. In *2016 International conference on big data and smart computing (BigComp)*, pages 239–244. IEEE, 2016.
- [10] J. H. Ahn, J. S. Lee, C. Baek, and Y. Kim. Performance improvement of a dehumidifying heat pump using an additional waste heat source in electric vehicles with low occupancy. *Energy*, 115:67–75, 2016.
- [11] G. Akar, K. J. Clifton, and S. T. Doherty. Discretionary activity location choice: In-home or out-of-home? *Transportation*, 38:101–122, 2011.
- [12] M. Akhtar and S. Moridpour. A review of traffic congestion prediction using artificial intelligence. *Journal of Advanced Transportation*, pages 1–18, 2021.
- [13] Abdulrahman Al-Molegi, Mohammed Jabreel, and Baraq Ghaleb. Stf-rnn: Space time features-based recurrent neural network for predicting people next location. In *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1–7. IEEE, 2016.
- [14] M. Alexander. Modeling charging demand for 2030. Integrated Energy Policy Report, Retrieved from: <https://www.energy.ca.gov>, 2020.
- [15] J. Auld, O. Verbas, M. Javanmardi, and A. Rousseau. Impact of privately-owned level 4 cav technologies on travel demand and energy. *Procedia Computer Science*, 130:914–919, 2018.
- [16] N. Lakshitha D. Haputhanthri B. Dissanayake, O. Hemachandra and A. Wijayasiri. A comparison of arimax, var and lstm on multivariate short-term traffic volume forecasting. In *Conference of Open Innovations Association, FRUCT*, number 28, pages 564–570, 2021.

- [17] P. Bansal and K. M. Kockelman. Are we ready to embrace connected and self-driving vehicles? a case study of texans. *Transportation*, 45(2):641–675, 2018.
- [18] Hugo Barbosa, Marc Barthelemy, Gourab Ghoshal, Charlotte R James, Maxime Lenormand, Thomas Louail, Ronaldo Menezes, José J Ramasco, Filippo Simini, and Marcello Tomasini. Human mobility: Models and applications. *Physics Reports*, 734:1–74, 2018.
- [19] B. Barrett, B. Ran, and R. Pillai. Developing a dynamic traffic management modeling framework for hurricane evacuation. *Transportation Research Record*, 1733(1):115–121, 2000.
- [20] M. J. Beck and D. A. Hensher. Insights into the impact of covid-19 on household travel and activities in australia—the early days under restrictions. *Transport Policy*, 96:76–93, 2020.
- [21] A. Bedir, N. Crisostomo, J. Allen, E. Wood, and C. Rames. California plug-in electric vehicle infrastructure projections: 2017-2025. Technical Report CEC-600-2018-1, California Energy Commission, 2018.
- [22] Moshe E Ben-Akiva and John L Bowman. Activity based travel demand model systems. In *Equilibrium and advanced transportation modelling*, pages 27–46. Springer, 1998.
- [23] E. Ben-Elia, B. Alexander, C. Hubers, and D. Ettema. Activity fragmentation, ict and travel: An exploratory path analysis of spatiotemporal interrelationships. *Transportation Research Part A: Policy and Practice*, 68:56–74, 2014.
- [24] Chandra R Bhat, Konstadinos G Goulias, Ram M Pendyala, Rajesh Paleti, Raghuprasad Sidharthan, Laura Schmitt, and Hsi-hwa Hu. A household-level activity pattern generation model for the simulator of activities, greenhouse emissions, networks, and travel

- (simagent) system in southern california. In *91st Annual Meeting of the Transportation Research Board, Washington, DC*, 2012.
- [25] M. Bierlaire. A short introduction to biogeme. Technical report, Transport and Mobility Laboratory, ENAC, EPFL, 2023. Technical report TRANSP-OR 230620.
- [26] D. Billings and J. S. Yang. Application of the arima models to urban roadway travel time prediction—a case study. In *2006 IEEE International Conference on Systems, Man and Cybernetics*, volume 3, pages 2529–2534. IEEE, 2006.
- [27] California Air Resource Board. The clean cars 4 all program. Retrieved from <https://ww2.arb.ca.gov/>, 2021.
- [28] California Air Resource Board. Emission factor. Retrieved from: <https://arb.ca.gov>, 2022.
- [29] T. Bogaerts, A. D. Masegosa, J. S. Angarita-Zapata, E. Onieva, and P. Hellinckx. A graph cnn-lstm neural network for short and long-term traffic forecasting based on trajectory data. *Transportation Research Part C: Emerging Technologies*, 112:62–77, 2020.
- [30] M Bohm, Mirco Nanni, and Luca Pappalardo. Quantifying the presence of air pollutants over a road network in high spatio-temporal resolution. In *Climate Change AI, NeurIPS Workshop*, 2021.
- [31] John L Bowman and Moshe E Ben-Akiva. Activity-based disaggregate travel demand model system with activity schedules. *Transportation research part a: policy and practice*, 35(1):1–28, 2001.
- [32] Mark Bradley and John L Bowman. Design features of activity-based microsimulation models for us metropolitan planning organizations: a summary. In *Transportation Research Board Conference Proceedings*, volume 2, 2008.

- [33] R. P. Brooker and N. Qin. Identification of potential locations of electric vehicle supply equipment. *Journal of Power Sources*, 299:76–84, 2015.
- [34] K. E. Brown and R. Dodder. Energy and emissions implications of automated vehicles in the us energy system. *Transportation Research Part D: Transport and Environment*, 77:132–147, 2019.
- [35] US Census Bureau. Quickfacts. Retrieved from: <https://www.census.gov>, 2022.
- [36] B. Caldarola and S. Sorrell. Do teleworkers travel less? evidence from the english national travel survey. *Transportation Research Part A: Policy and Practice*, 159:282–303, 2022.
- [37] California Department of Transportation. PeMS, 2020. Retrieved from California Department of Transportation.
- [38] E. Canal, M. Ward, R. Knight, and S. Lebowitz. These companies are making workers return to the office, despite the rise of remote work. Retrieved from <https://www.businessinsider.com/>, 2023.
- [39] Jinzhou Cao, Qingquan Li, Wei Tu, and Feilong Wang. Characterizing preferred motif choices and distance impacts. *Plos one*, 14(4):e0215242, 2019.
- [40] X. Cao, Z. Xu, and F. Douma. The interactions between e-shopping and traditional in-store shopping: An application of structural equations model. *Transportation*, 39(5):957–974, 2012.
- [41] Ennio Cascetta. *Transportation systems analysis: models and applications*, volume 29. Springer Science & Business Media, 2009.
- [42] National Hurricane Center and Central Pacific Hurricane Center. National hurricane center and central pacific hurricane center.

- [43] S. Chakrabarti. Does telecommuting promote sustainable travel and physical activity? *Journal of Transport and Health*, 9:19–33, 2018.
- [44] P. Chatterjee, P. Edara, S. Menneni, and C. Sun. Replication of work zone capacity values in a simulation model. *Transportation Research Record*, 2130(1):138–148, 2009.
- [45] S. Childress, B. Nichols, B. Charlton, and S. Coe. Using an activity-based model to explore the potential impacts of automated vehicles. *Transportation Research Record*, 2493(1):99–106, 2015.
- [46] J. Y. Chow, K. Ozbay, Y. He, J. Zhou, M. Lee, D. Wang, and D. Sha. Multi-agent simulation-based virtual test bed ecosystem: Matsim-nyc. Technical report, 2020.
- [47] L. Chu, S. Oh, and W. Recker. Adaptive kalman filter based freeway travel time estimation. In *84th TRB Annual Meeting, Washington DC*, 2005.
- [48] G. Circella, M. Jaller, R. Sun, X. Qian, and F. Alemi. Emissions impact of connected and automated vehicle deployment in california. Technical report, 2021.
- [49] P. Coate. Remote work before, during, and after the pandemic, 2021.
- [50] J. Cochran, P. Denholm, M. Mooney, D. Steinberg, E. Hale, G. Heath, and S. Nicholson. La100: The los angeles 100 Technical report, National Renewable Energy Laboratory, 2021.
- [51] Carmela Comito. Next: A framework for next-place prediction on location based social networks. *Knowledge-Based Systems*, 204:106205, 2020.
- [52] Joint Technology Committee. Teleservices report, 2019.
- [53] US Congress. Telework enhancement act of 2010, 2010.
- [54] H. Couclelis. Rethinking time geography in the information age. *Environment and Planning A*, 41:1556–1575, 2009.

- [55] Puget Sound Regional Council. Puget sound regional council household travel survey. <https://www.nrel.gov>, 2017.
- [56] H. Cui, F. Li, X. Fang, and R. Long. Distribution network reconfiguration with aggregated electric vehicle charging strategy. In *2015 IEEE Power & Energy Society General Meeting*, pages 1–5. IEEE, 2015.
- [57] H. Jeon D. Jo, B. Yu and K. Sohn. Image-to-image learning to predict traffic speeds by considering area-wide spatio-temporal dependencies. *IEEE Transactions on Vehicular Technology*, 68(2):1188–1197, 2018.
- [58] S. Dahal, J. Hernandez, and J. Roesler. Infrastructure enhancements for cav navigation (no. ict-20-008, uilu-eng-2020-2008). Technical report, University of Michigan. Center for Connected and Automated Transportation, 2018.
- [59] N. Daina, A. Sivakumar, and J. W. Polak. Modelling electric vehicles use: a survey on the methods. *Renewable and Sustainable Energy Reviews*, 68:447–460, 2017.
- [60] M. Dave. The crucial reason houston officials didn’t order evacuations before harvey made landfall.
- [61] R. A. Davidson, L. K. Nozick, T. Wachtendorf, B. Blanton, B. Colle, R. L. Kolar, and N. Leonardo. An integrated scenario ensemble-based framework for hurricane evacuation modeling: Part 1—decision support system. *Risk analysis*, 40(1):97–116, 2020.
- [62] J. De Abreu e Silva and P. C. Melo. Home telework, travel behavior, and land-use patterns. *Journal of Transport and Land Use*, 11(1):419–441, 2018.
- [63] Deloitte. Deloitte city mobility index 2020. Retrieved from: [www.deloitte.com](http://www.deloitte.com), 2020.
- [64] M. A. Delucchi and T. E. Lipman. *Lifetime cost of battery, fuel-cell, and plug-in hybrid electric vehicles*, pages 19–60. Amsterdam, The Netherlands, 2010.



- [65] Jean-Claude Deville, Carl-Erik Särndal, and Olivier Sautory. Generalized raking procedures in survey sampling. *Journal of the American statistical Association*, 88(423):1013–1020, 1993.
- [66] V. Dixit, T. Montz, and B. Wolshon. Validation techniques for region-level microscopic mass evacuation traffic simulations. *Transportation research record*, 2229(1):66–74, 2011.
- [67] J. Dong, C. Liu, and Z. Lin. Charging infrastructure planning for promoting battery electric vehicles: An activity-based approach using multiday travel data. *Transportation Research Part C: Emerging Technologies*, 38:44–55, 2014.
- [68] Y. Duan, L. V. Yisheng, and F. Y. Wang. Travel time prediction with lstm neural network. In *2016 IEEE 19th international conference on intelligent transportation systems (ITSC)*, pages 1053–1058. IEEE, 2016.
- [69] C. Duffy. Big tech firms ramp up remote working orders to prevent coronavirus spread. Retrieved from <https://www.cnn.com>, 2020.
- [70] Clare Duffy. Big tech firms ramp up remote working orders to prevent coronavirus spread. <https://www.cnn.com/>, 2020. Accessed: 2024-05-15.
- [71] Dynata. Dynata data services. Retrieved from: <https://www.dynata.com>.
- [72] P. Edara, S. Sharma, and C. McGhee. Development of a large-scale traffic simulation model for hurricane evacuation—methodology and lessons learned. *Natural Hazards Review*, 11(4):127–139, 2010.
- [73] D. J. Fagnant and K. M. Kockelman. The travel and environmental implications of shared autonomous vehicles, using agent-based model scenarios. *Transportation Research Part C: Emerging Technologies*, 40:1–13, 2014.

- [74] Federal Highway Administration. 2022 nextgen national household travel survey core data, 2022.
- [75] Federal Highway Administration (FHWA). Making work zones work better. <https://ops.fhwa.dot.gov>, 2004.
- [76] Federal Highway Administration (FHWA). Using modeling and simulation tools for work zone analysis. <https://ops.fhwa.dot.gov/>, 2009.
- [77] Federal Highway Administration (FHWA). Work zone data exchange (wzdx). <https://ops.fhwa.dot.gov>, 2022.
- [78] Federal Highway Administration (FHWA). Fhwa work zone facts and statistics. <https://ops.fhwa.dot.gov>, 2024.
- [79] Jie Feng, Yong Li, Chao Zhang, Funing Sun, Fanchao Meng, Ang Guo, and Depeng Jin. Deepmove: Predicting human mobility with attentional recurrent networks. In *Proceedings of the 2018 world wide web conference*, pages 1459–1468, 2018.
- [80] K. Feng and N. Lin. Modeling and analyzing the traffic flow during evacuation in hurricane irma (2017). *Transportation Research Part D: Transport and Environment*, 110:103412, 2022.
- [81] G. M. Fetene, S. Kaplan, S. L. Mabit, A. F. Jensen, and C. G. Prato. Harnessing big data for estimating the energy consumption and driving range of electric vehicles. *Transportation Research Part D: Transport and Environment*, 54:1–11, 2017.
- [82] A. Foley, B. Tyther, P. Calnan, and B. Ó. Gallachóir. Impacts of electric vehicle charging under electricity market operations. *Applied Energy*, 101:93–102, 2013.
- [83] K. T. Geurs and B. Van Wee. Accessibility evaluation of land-use and transport strategies: review and research directions. *Journal of Transport Geography*, 12(2):127–140, 2004.

- [84] Z. Ghandeharioun and A. Kouvelas. Link travel time estimation for arterial networks based on sparse gps data and considering progressive correlations. *IEEE Open Journal of Intelligent Transportation Systems*, 3:679–694, 2022.
- [85] M. Gharbaoui, B. Martini, R. Bruno, L. Valcarenghi, M. Conti, and P. Castoldi. Designing and evaluating activity-based electric vehicle charging in urban areas. In *2013 IEEE International Electric Vehicle Conference (IEVC)*, pages 1–5. IEEE, 2013.
- [86] M. Ghorbanzadeh, S. Burns, L. V. N. Rugminiamma, E. Erman Ozguven, and W. Huang. Spatiotemporal analysis of highway traffic patterns in hurricane irma evacuation. *Transportation research record*, 2675(9):321–334, 2021.
- [87] E. Giovanis. The relationship between teleworking, traffic and air pollution. *Atmospheric Pollution Research*, 9(1):1–14, 2018.
- [88] Marta C Gonzalez, Cesar A Hidalgo, and Albert-Laszlo Barabasi. Understanding individual human mobility patterns. *nature*, 453(7196):779–782, 2008.
- [89] Konstadinos G Goulias, Chandra R Bhat, Ram M Pendyala, Yali Chen, Rajesh Paleti, Karthik C Konduri, Guoxiong Huang, and Hsi-Hwa Hu. Simulator of activities, greenhouse emissions, networks, and travel (simagent) in southern california: Design, implementation, preliminary findings, and integration plans. In *2011 IEEE Forum on Integrated and Sustainable Transportation Systems*, pages 164–169. IEEE, 2011.
- [90] The Guardian. Hurricane ida: thousands evacuate from new orleans as storm bears down, 2021.
- [91] M. Gucwa. *Mobility and energy impacts of automated cars*. PhD thesis, Stanford University, 2014.

- [92] Xiaogang Guo, Zhijie Xu, Jianqin Zhang, Jian Lu, and Hao Zhang. An od flow clustering method based on vector constraints: a case study for beijing taxi origin-destination data. *ISPRS International Journal of Geo-Information*, 9(2):128, 2020.
- [93] S. Habib, M. M. Khan, F. Abbas, M. Numan, Y. Ali, H. Tang, and X. Yan. A framework for stochastic estimation of electric vehicle charging behavior for risk assessment of distribution networks. *Frontiers in Energy*, 14(2):298–317, 2020.
- [94] F. G. Habtemichael and M. Cetin. Short-term traffic flow rate forecasting based on identifying similar traffic patterns. *Transportation research Part C: emerging technologies*, 66:61–78, 2016.
- [95] S. Hardman, A. Jenn, G. Tal, J. Axsen, G. Beard, N. Daina, and B. Witkamp. A review of consumer preferences of and interactions with electric vehicle charging infrastructure. *Transportation Research Part D: Transport and Environment*, 62:508–523, 2018.
- [96] S. Hasan and G. Foliente. Modeling infrastructure system interdependencies and socioeconomic impacts of failure in extreme events: emerging r&d challenges. *Natural Hazards*, 78(3):2143–2168, 2015.
- [97] S. Hasan, R. Mesa-Arango, and S. Ukkusuri. A random-parameter hazard-based model to understand household evacuation timing behavior. *Transportation research part C: emerging technologies*, 27:108–116, 2013.
- [98] S. Hasan, S. Ukkusuri, H. Gladwin, and P. Murray-Tuite. Behavioral model to understand household-level hurricane evacuation decision making. *Journal of Transportation Engineering*, 137(5):341–348, 2011.
- [99] B. Y. He, Q. Jiang, and J. Ma. Connected automated vehicle impacts in southern california part-i: travel behavior and demand analysis. *Transportation Research Part D: Transport and Environment*, 2022. under review.

- [100] B. Y. He, Q. Jiang, and J. Ma. Connected automated vehicle impacts in southern california part-i: Travel behavior and demand analysis. *Transportation Research Part D: Transport and Environment*, 109:103329, 2022.
- [101] B. Y. He, J. Zhou, Z. Ma, D. Wang, D. Sha, M. Lee, J. Y. Chow, and K. Ozbay. A validated multi-agent simulation test bed to evaluate congestion pricing policies on population segments by time-of-day in new york city. *Transport Policy*, 101:145–161, 2021.
- [102] B. Y. He, J. Zhou, Z. Ma, D. Wang, D. Sha, M. Lee, and K. Ozbay. A validated multi-agent simulation test bed to evaluate congestion pricing policies on population segments by time-of-day in new york city. *Transport Policy*, 101:145–161, 2021.
- [103] Brian Yueshuai He, Qinhua Jiang, Ma Haoxuan, and Jiaqi Ma. Multi-agent multimodal transportation simulation for mega-cities: Application of los angeles. *Procedia Computer Science*, 238:736–741, 2024.
- [104] Brian Yueshuai He, Qinhua Jiang, and Jiaqi Ma. Connected automated vehicle impacts in southern california part-i: Travel behavior and demand analysis. *Transportation research part D: transport and environment*, 109:103329, 2022.
- [105] A. Henao and W. E. Marshall. The impact of ride-hailing on vehicle miles traveled. *Transportation*, 46(6):2173–2194, 2019.
- [106] David A. Hensher and Peter R. Stopher. Behavioural travel modelling. In *Behavioural travel modelling*, pages 11–52. Routledge, 2021.
- [107] A. D. Hilshey, P. D. Hines, P. Rezaei, and J. R. Dowds. Estimating the impact of electric vehicle smart charging on distribution transformer aging. *IEEE Transactions on Smart Grid*, 4(2):905–913, 2012.

- [108] A. Horni, K. Nagel, and K. W. Axhausen. *Introducing MATSim*, page 3–8. Ubiquity Press, 2016.
- [109] Y. Hou, P. Edara, and C. Sun. Traffic flow forecasting for urban work zones. *IEEE Transactions on Intelligent Transportation Systems*, 16(4):1761–1770, 2014.
- [110] C. W. Hsu and K. Fingerman. Public electric vehicle charger access disparities across race and income in california. *Transport Policy*, 100:59–67, 2021.
- [111] Y. Huang, K. M. Kockelman, V. Garikapati, L. Zhu, and S. Young. Use of shared automated vehicles for first-mile last-mile service: micro-simulation of rail-transit connections in austin, texas. *Transportation Research Record*, 2675(2):135–149, 2021.
- [112] C. Ihrig. Travel cost savings and practicality for low-vision telerehabilitation. *Telemedicine and E-Health*, 25(7):649–654, 2019.
- [113] Fortune Business Insights. Us electric vehicle market size, share & covid-19 impact analysis, by vehicle type and regional forecast, 2021-2028. Retrieved from: <https://www.fortunebusinessinsights.com>, 2020.
- [114] McKinsey Global Institute. What’s next for remote work: An analysis of 2,000 tasks, 800 jobs, and nine countries. Retrieved from: <https://www.mckinsey.com>, 2020.
- [115] K. Choi J. Bae and J. H. Oh. Multicontextual machine learning approach to modeling traffic impact of urban highway work zones. *Transportation Research Record: Journal of the Transportation Research Board*, 2645:184–194, 2017.
- [116] J. K. Mathew H. Li J. Desai, B. Scholer and D. M. Bullock. Analysis of route choice during planned and unplanned road closures. *IEEE Open Journal of Intelligent Transportation Systems*, 3:489–502, 2022.

- [117] Q. Jiang, B. Y. He, and J. Ma. Connected automated vehicle impacts in southern california part-ii: Vmt, emissions, and equity. *Transportation Research Part D: Transport and Environment*, 109:103381, 2022.
- [118] Q. Jiang, B. Y. He, and J. Ma. Connected automated vehicle impacts in southern california part-ii: Vmt, emissions, and equity. *Transportation Research Part D: Transport and Environment*, 109:103381, 2022.
- [119] Qinhuia Jiang, Brian Yueshuai He, Changju Lee, and Jiaqi Ma. Deploying scalable traffic prediction models for efficient management in real-world large transportation networks during hurricane evacuations. *arXiv preprint arXiv:2406.12119*, 2024.
- [120] Qinhuia Jiang, Brian Yueshuai He, and Jiaqi Ma. Connected automated vehicle impacts in southern california part-ii: Vmt, emissions, and equity. *Transportation research part D: transport and environment*, 109:103381, 2022.
- [121] Qinhuia Jiang, Xishun Liao, Yaofa Gong, and Jiaqi Ma. An attention-based multi-context convolutional encoder-decoder neural network for work zone traffic impact prediction. *arXiv preprint arXiv:2405.21045*, 2024.
- [122] Qinhuia Jiang, Ning Zhang, Brian Yueshuai He, Changju Lee, and Jiaqi Ma. Large-scale public charging demand prediction with a scenario-and activity-based approach. *Transportation research part A: Policy and Practice*, 179:103935, 2024.
- [123] Renhe Jiang, Xuan Song, Zipei Fan, Tianqi Xia, Qunjun Chen, Satoshi Miyazawa, and Ryosuke Shibasaki. Deepurbanmomentum: An online deep-learning system for short-term urban mobility prediction. In *Proceedings of the AAAI conference on artificial intelligence*, volume 32, 2018.
- [124] Shan Jiang, Yingxiang Yang, Siddharth Gupta, Daniele Veneziano, Shounak Athavale, and Marta C González. The timegeo modeling framework for urban mobility without

- travel surveys. *Proceedings of the National Academy of Sciences*, 113(37):E5370–E5378, 2016.
- [125] Raja Jurdak, Kun Zhao, Jiajun Liu, Maurice AbouJaoude, Mark Cameron, and David Newth. Understanding human mobility from twitter. *PloS one*, 10(7):e0131469, 2015.
- [126] G. Kalton and I. Flores-Cervantes. Weighting methods. *Journal of Official Statistics*, 19(2):81–97, 2003.
- [127] Dmytro Karamshuk, Chiara Boldrini, Marco Conti, and Andrea Passarella. Human mobility models for opportunistic networks. *IEEE Communications Magazine*, 49(12):157–165, 2011.
- [128] Karim and H. Adeli. Radial basis function neural network for work zone capacity and queue estimation. *Journal of Transportation Engineering*, 129(5):494–503, 2003.
- [129] M. Kaviani-pour, F. Fakharmoosavi, H. Singh, M. Ghamami, A. Zockaie, Y. Ouyang, and R. Jackson. Electric vehicle fast charging infrastructure planning in urban networks considering daily travel and charging behavior. *Transportation Research Part D: Transport and Environment*, 93:102769, 2021.
- [130] S. M. Khan, M. Chowdhury, E. A. Morris, and L. Deka. Synergizing roadway infrastructure investment with digital infrastructure for infrastructure-based connected vehicle applications: Review of current status and future directions. *Journal of Infrastructure Systems*, 25(4):03119001, 2019.
- [131] S. N. Kim. Is telecommuting sustainable? an alternative approach to estimating the impact of home-based telecommuting on household travel. *International Journal of Sustainable Transportation*, 11:72–85, 2017.
- [132] S.-N. Kim, S. Choo, and P. L. Mokhtarian. Home-based telecommuting and intra-household interactions in work and non-work travel: A seemingly unrelated censored



- regression approach. *Transportation Research Part A: Policy and Practice*, 80:197–214, 2015.
- [133] L. Knapen, B. Kochan, T. Bellemans, D. Janssens, and G. Wets. Activity based models for countrywide electric vehicle power demand calculation. In *2011 IEEE First International Workshop on Smart Grid Modeling and Simulation (SGMS)*, pages 13–18. IEEE, 2011.
- [134] P. Koopman and M. Wagner. Toward a framework for highly automated vehicle safety validation. Technical report, 2018. SAE Technical Paper, Tech. Rep.
- [135] Moritz UG Kraemer, Chia-Hung Yang, Bernardo Gutierrez, Chieh-Hsi Wu, Brennan Klein, David M Pigott, Open COVID-19 Data Working Group†, Louis Du Plessis, Nuno R Faria, Ruoran Li, et al. The effect of human mobility and control measures on the covid-19 epidemic in china. *Science*, 368(6490):493–497, 2020.
- [136] Kundan Krishna, Deepali Jain, Sanket V Mehta, and Sunav Choudhary. An lstm based system for prediction of human activities with durations. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(4):1–31, 2018.
- [137] A. Lari. Telework/workforce flexibility to reduce congestion and environmental degradation? *Procedia-Social and Behavioral Sciences*, 48:712–721, 2012.
- [138] Tousignant Lauren. The cost of natural disasters nearly doubled in 2017, 2017.
- [139] Juha K Laurila, Daniel Gatica-Perez, Imad Aad, Olivier Bornet, Trinh-Minh-Tri Do, Olivier Dousse, Julien Eberle, Markus Miettinen, et al. The mobile data challenge: Big data for mobile computing research. In *Pervasive computing*, 2012.
- [140] S. Le Vine, C. Latinopoulos, and J. Polak. Analysis of the relationship between internet usage and allocation of time for personal travel and out of home activities: Case study of scotland in 2005/6. *Travel Behaviour and Society*, 4:49–59, 2016.

- [141] R. Lee, S. Yazbeck, and S. Brown. Validation and application of agent-based electric vehicle charging model. *Energy Reports*, 6:53–62, 2020.
- [142] R. C. Leou, J. H. Teng, H. J. Lu, B. R. Lan, H. T. Chen, T. Y. Hsieh, and C. L. Su. Stochastic analysis of electric transportation charging impacts on power quality of distribution systems. *IET Generation, Transmission & Distribution*, 12(11):2725–2734, 2018.
- [143] D. Levinson. Equity effects of road pricing: A review. *Transport Reviews*, 30(1):33–57, 2010.
- [144] Linchao Li, Xu Qu, Jian Zhang, Yonggang Wang, and Bin Ran. Traffic speed prediction for intelligent transportation system based on a deep feature fusion model. *Journal of Intelligent Transportation Systems*, 23(6):605–616, 2019.
- [145] S. Li. Revisiting the relationship between information and communication technologies and travel behavior: An investigation of older americans. *Transportation Research Part A: Policy and Practice*, 172:103689, 2023.
- [146] Dongliang Liao, Weiqing Liu, Yuan Zhong, Jing Li, and Guowei Wang. Predicting activity and location with multi-task context aware recurrent neural network. In *IJCAI*, pages 3435–3441, 2018.
- [147] Xishun Liao, Brian Yueshuai He, Qinhua Jiang, Chenchen Kuai, and Jiaqi Ma. Deep activity model: A generative approach for human mobility pattern synthesis. *arXiv preprint arXiv:2405.17468*, 2024.
- [148] M. K. Lindell, P. Murray-Tuite, B. Wolshon, and E. J. Baker. *Large-scale evacuation: The analysis, modeling, and management of emergency relocation from hazardous areas*. CRC Press, 2018.

- [149] J. Liu, K. M. Kockelman, P. M. Boesch, and F. Ciari. Tracking a system of shared autonomous vehicles across the austin, texas network using agent-based simulation. *Transportation*, 44(6):1261–1278, 2017.
- [150] J. Liu, J. Lipsitt, M. Jerrett, and Y. Zhu. Decreases in near-road no and no2 concentrations during the covid-19 pandemic in california. *Environmental Science & Technology Letters*, 8(2):161–167, 2020.
- [151] Qiang Liu, Shu Wu, Liang Wang, and Tieniu Tan. Predicting the next location: A recurrent model with spatial and temporal contexts. In *Proceedings of the AAAI conference on artificial intelligence*, volume 30, 2016.
- [152] R. Liu, Z. Ding, X. Jiang, J. Sun, Y. Jiang, and W. Qiang. How does experience impact the adoption willingness of battery electric vehicles? the role of psychological factors. *Environmental Science and Pollution Research*, 27(20):25230–25247, 2020.
- [153] X. Liu and Z. Bie. Optimal allocation planning for public ev charging station considering ac and dc integrated chargers. *Energy Procedia*, 159:382–387, 2019.
- [154] Yifan Liu, Chenchen Kuai, Haoxuan Ma, Xishun Liao, Brian Yueshuai He, and Jiaqi Ma. Semantic trajectory data mining with llm-informed poi classification, 2024.
- [155] Yifan Liu, Xishun Liao, Haoxuan Ma, Brian Yueshuai He, Chris Stanford, and Jiaqi Ma. Human mobility modeling with limited information via large language models. *arXiv preprint arXiv:2409.17495*, 2024.
- [156] Massimiliano Luca, Gianni Barlacchi, Bruno Lepri, and Luca Pappalardo. A survey on deep learning for human mobility. *ACM Computing Surveys (CSUR)*, 55(1):1–44, 2021.
- [157] S. Choudhury M. Emu, F. B. Kamal and Q. A. Rahman. Fatality prediction for motor vehicle collisions: Mining big data using deep learning and ensemble methods. *IEEE Open Journal of Intelligent Transportation Systems*, 3:199–209, 2022.

- [158] H. Wu Q. Huang M. Tang, X. Fu and Q. Zhao. Traffic flow anomaly detection based on robust ridge regression with particle swarm optimization algorithm. *Mathematical Problems in Engineering*, pages 1–10, 2020.
- [159] J. Kim M. Zhong and Z. Zheng. Estimating link flows in road networks with synthetic trajectory data generation: Inverse reinforcement learning approach. In *IEEE Open Journal of Intelligent Transportation Systems*, volume 4, pages 14–29, 2023.
- [160] J. Ma, J. Hu, E. Leslie, F. Zhou, P. Huang, and J. Bared. An eco-drive experiment on rolling terrains for fuel consumption optimization with connected automated vehicles. *Transportation Research Part C: Emerging Technologies*, 100:125–141, 2019.
- [161] Jiaqi Ma, Yueshuai He, Qinhua Jiang, METRANS Transportation Center, et al. Evaluating policies and incentives to reduce vehicle-miles-traveled and air pollutant emissions through the promotion of telework and remote services. Technical report, METRANS Transportation Center (Calif.), 2022.
- [162] H. S. Mahmassani, J. Kim, Y. Chen, Y. Stogios, A. Brijmohan, and P. Vovsha. Incorporating reliability performance measures into operations and planning modeling tools. In *Transportation Research Board Annual Meeting*, 2014.
- [163] Nancy McGuckin, Anthony Fucci, et al. Summary of travel trends: 2017 national household travel survey. Technical report, United States. Department of Transportation. Federal Highway Administration, 2018.
- [164] L. Mendelson. Stay on top of "stay at home" – a list of statewide orders. Retrieved from: <https://www.littler.com/>, 2020.
- [165] Q. Meng and J. Weng. Evaluation of rear-end crash risk at work zone using work zone traffic data. *Accident Analysis & Prevention*, 43(4):1291–1300, 2011.

- [166] R. Mesa-Arango, S. Hasan, S. V. Ukkusuri, and P. Murray-Tuite. Household-level model for hurricane evacuation destination type choice using hurricane ivan data. *Natural hazards review*, 14(1):11–20, 2013.
- [167] Metro. I-405 comprehensive multimodal corridor plan. Retrieved from: <https://www.metro.net>, 2022.
- [168] D. Milakis, M. Kroesen, and B. van Wee. Implications of automated vehicles for accessibility and location choices: Evidence from an expert-based experiment. *Journal of Transport Geography*, 68:142–148, 2018.
- [169] D. Milakis, B. Van Arem, and B. Van Wee. Policy and society related implications of automated driving: A review of literature and directions for future research. *Journal of Intelligent Transportation Systems*, 21(4):324–348, 2017.
- [170] H. Mitomo and H. Oniki. Information technology for sustainable societies—public policy perspectives in japan: the case of telework. *The IPTS Report*, 32:24–31, 1999.
- [171] Neda Mohammadi and John E Taylor. Urban energy flux: Spatiotemporal fluctuations of building energy consumption and human mobility-driven prediction. *Applied energy*, 195:810–818, 2017.
- [172] P. L. Mokhtarian. A synthetic approach to estimating the impacts of telecommuting on travel. *Urban Studies*, 35(2):215–241, 1998.
- [173] P. L. Mokhtarian. Telecommunications and travel: The case for complementarity. *Journal of Industrial Ecology*, 6:43–57, 2003.
- [174] P. L. Mokhtarian. If telecommunication is such a good substitute for travel, why does congestion continue to get worse? *Transportation Letters*, 1:1–17, 2009.
- [175] Dietmar P.F. Möller. *Introduction to transportation analysis, modeling and simulation*. Simulation Foundations, Methods and Applications. Springer London, 2014.

- [176] T. Montz and Z. Zhang. Modeling regional hurricane evacuation events: calibration and validation. *Natural Hazards Review*, 16(4):04015007, 2015.
- [177] J. Mullan, D. Harries, T. Bräunl, and S. Whitely. Modelling the impacts of electric vehicle recharging on the western australian electricity supply system. *Energy Policy*, 39(7):4349–4359, 2011.
- [178] Werabhat Mungthanya, Santi Phithakkitnukoon, Merkebe Getachew Demissie, Lina Kattan, Marco Veloso, Carlos Bento, and Carlo Ratti. Constructing time-dependent origin-destination matrices with adaptive zoning scheme and measuring their similarities with taxi trajectory data. *IEEE Access*, 7:77723–77737, 2019.
- [179] P. Murray-Tuite and B. Wolshon. Evacuation transportation modeling: An overview of research, development, and practice. *Transportation Research Part C: Emerging Technologies*, 27:25–45, 2013.
- [180] J. Myung, D. K. Kim, S. Y. Kho, and C. H. Park. Travel time prediction using k nearest neighbor method with combined data from vehicle detector system and automatic toll collection system. *Transportation Research Record*, 2256(1):51–59, 2011.
- [181] H. P. Zhao H. Kim N. Ranjan, S. Bhandari and P. Khan. City-wide traffic congestion prediction based on cnn, lstm and transpose cnn. *IEEE Access*, 8:81606–81620, 2020.
- [182] B. H. Nahmias-Biran, J. B. Oke, and N. Kumar. Who benefits from avs? equity implications of automated vehicles policies in full-scale prototype cities. *Transportation Research Part A: Policy and Practice*, 154:92–107, 2021.
- [183] E. Narassimhan and C. Johnson. The role of demand-side incentives and charging infrastructure on plug-in electric vehicle adoption: analysis of us states. *Environmental Research Letters*, 13(7):074032, 2018.

- [184] National Renewable Energy Laboratory. Transportation Secure Data Center. Accessed Jan. 15, 2019, 2019.
- [185] A. Neumann. *Berlin I: BVG scenario*, page 369–370. Ubiquity Press, 2016.
- [186] New York Metropolitan Transportation Council. 2010-2011 regional household travel survey (rhst), 2012. Retrieved from: <https://www.nymtc.org>.
- [187] E. Newburger. Covid pandemic drove a record drop in global carbon emissions in 2020. CNBC, Retrieved from: <https://www.cnbc.com/>, 2020.
- [188] National Highway Traffic Safety Administration (NHTSA). 2016 motor vehicle crashes: overview. traffic safety facts: research note. Technical report, 2017.
- [189] D. Ni, J. Li, S. Andrews, and H. Wang. A methodology to estimate capacity impact due to connected vehicle technology. *International Journal of Vehicular Technology*, 2012.
- [190] M. Nour, A. Ali, and C. Farkas. Evaluation of electric vehicles charging impacts on a real low voltage grid. *International Journal of Power Engineering and Energy*, 9(2):837–842, 2018.
- [191] NREL. Cleansed data from household travel studies and surveys, 2023. Retrieved from: <https://www.nrel.gov>.
- [192] L. L. C. NuStats. 2010–2012 california household travel survey final report. Technical report, 2013.
- [193] National Academy of Engineering Committee on Telecommunications. Telecommunications for enhanced metropolitan function and form: Report to the director of telecommunications management. Technical report, National Academy of Engineering, 1969.

- [194] Southern California Association of Governments (SCAG). Scag transportation demand management strategic plan and final report. Technical report, 2019.
- [195] Southern California Association of Governments (SCAG). 2016 regional travel demand model and model validation report. Technical report, 2020.
- [196] Bureau of Labor Statistics. American time use survey — 2019 results. Retrieved from <https://www.bls.gov>, 2020.
- [197] National Institute of Statistics and Geography. Origin-destination survey in households of the metropolitan zone of the valley of mexico. <https://en.www.inegi.org.mx>, 2017.
- [198] California Department of Transportation. California household travel survey. <https://www.nrel.gov/transportation>, 2012.
- [199] U.S. Department of Transportation (US DOT). Automated vehicle research. Retrieved from <https://www.its.dot.gov/>.
- [200] S. Oh, Y. J. Byon, K. Jang, and H. Yeo. Short-term travel-time prediction on highway: A review on model-based approach. *KSCE Journal of Civil Engineering*, 22:298–310, 2018.
- [201] A. Olia, S. Razavi, B. Abdulhai, and H. Abdelgawad. Traffic capacity implications of automated vehicles mixed with regular vehicles. *Journal of Intelligent Transportation Systems*, 22(3):244–262, 2018.
- [202] A. Ostermann, Y. Fabel, and K. Ouan. Forecasting charging station occupancy using supervised learning algorithms, 2021.
- [203] Outsider. Hurricane ida causing major travel disruptions with canceled flights and massive traffic jams, 2021.



- [204] B. Ozbilen, K. Wang, and G. Akar. Revisiting the impacts of virtual mobility on travel behavior: An exploration of daily travel time expenditures. *Transportation Research Part A: Policy and Practice*, 145:49–62, 2021.
- [205] Shaobin Pan, Hongxia Yan, Jiefu He, and Zhiguo He. Vulnerability and resilience of transportation systems: A recent literature review. *Physica A: Statistical Mechanics and its Applications*, 581:126235, 2021.
- [206] I. Cardei Papa and M. Cardei. Generalized path planning for utm systems with a space-time graph. In *IEEE Open Journal of Intelligent Transportation Systems*, volume 3, pages 351–368, 2022.
- [207] Luca Pappalardo, Salvatore Rinzivillo, Zehui Qu, Dino Pedreschi, and Fosca Giannotti. Understanding the patterns of car travel. *The European Physical Journal Special Topics*, 215:61–73, 2013.
- [208] Luca Pappalardo and Filippo Simini. Data-driven generation of spatio-temporal routines in human mobility. *Data Mining and Knowledge Discovery*, 32(3):787–829, 2018.
- [209] Luca Pappalardo, Filippo Simini, Salvatore Rinzivillo, Dino Pedreschi, Fosca Giannotti, and Albert-László Barabási. Returners and explorers dichotomy in human mobility. *Nature communications*, 6(1):8166, 2015.
- [210] Luca Pappalardo, Maarten Vanhoof, Lorenzo Gabrielli, Zbigniew Smoreda, Dino Pedreschi, and Fosca Giannotti. An analytical framework to nowcast well-being using mobile phone data. *International Journal of Data Science and Analytics*, 2:75–92, 2016.
- [211] D. Park, L. R. Rilett, and G. Han. Spectral basis neural networks for real-time travel time forecasting. *Journal of Transportation Engineering*, 125(6):515–523, 1999.
- [212] K. Parker, J. M. Horowitz, and R. Minkin. How the coronavirus outbreak has—and hasn’t—changed the way americans work, 2020.

- [213] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.
- [214] A. L. Patwary and A. J. Khattak. Interaction between information and communication technologies and travel behavior: Using behavioral data to explore correlates of the covid-19 pandemic. *Transportation Research Record*, 2022.
- [215] California Performance Measurement System (PeMS). California performance measurement system. Retrieved from: <https://pems.dot.ca.gov>, 2020.
- [216] Marist Poll. Npr marist poll tables of questions may 2018. <https://maristpoll.marist.edu>, 2018.
- [217] A. Priessner, R. Sposato, and N. Hampl. Predictors of electric vehicle adoption: An analysis of potential electric vehicle drivers in austria. *Energy Policy*, 122:701–714, 2018.
- [218] J. Ma L. Rodegerdts B. Cesme A. Bibeka Q. Jiang, B. Schroeder and A. Morgan. Developing highway capacity manual capacity adjustment factors for connected and automated traffic on roundabouts. *Journal of Transportation Engineering, Part A: Systems*, 148(5):04022014, 2022.
- [219] Y. Guo M. Ahmed G. Yang Q. Jiang, D. Nian and J. Ma. Evaluating connected vehicle-based weather responsive management strategies using weather-sensitive microscopic simulation. *Journal of Intelligent Transportation Systems*, 27(1):92–110, 2023.
- [220] Z. Zhang R. Fu and L. Li. Using lstm and gru neural network methods for traffic flow prediction. In *2016 31st Youth academic annual conference of Chinese association of automation (YAC)*, pages 324–328. IEEE, 2016.

- [221] R. Rahman and S. Hasan. Short-term traffic speed prediction for freeways during hurricane evacuation: a deep learning approach. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 1291–1296. IEEE, 2018.
- [222] R. Rahman and S. Hasan. A deep learning approach for network-wide dynamic traffic prediction during hurricane evacuation. *Transportation Research Part C: Emerging Technologies*, 152:104126, 2023.
- [223] R. Rahman, S. Hasan, and M. H. Zaki. Towards reducing the number of crashes during hurricane evacuation: Assessing the potential safety impact of adaptive cruise control systems. *Transportation research part C: emerging technologies*, 128:103188, 2021.
- [224] T. Rambha, L. Nozick, and R. Davidson. Modeling departure time decisions during hurricanes using a dynamic discrete choice framework. In *No. 19-06045*, 2019.
- [225] N. Rieser-Schüssler, P. M. Bösch, A. Horni, and M. Balmer. *Zurich*, page 375–378. Ubiquity Press, 2016.
- [226] Salvatore Rinzivillo, Lorenzo Gabrielli, Mirco Nanni, Luca Pappalardo, Dino Pedreschi, and Fosca Giannotti. The purpose of motion: Learning activities from individual mobility networks. In *2014 International conference on data science and advanced analytics (DSAA)*, pages 312–318. IEEE, 2014.
- [227] Jesse Roberts. On the computational power of decoder-only transformer language models. *arXiv preprint arXiv:2305.17026*, 2023.
- [228] R. M. Robinson, P. Foytik, and C. Jordan. Review and analysis of user inputs to online evacuation modeling tool. *Transportation Research Board 96th Annual Meeting*, (17-06460), 2017.

- [229] K. C. Roy, S. Hasan, A. Culotta, and N. Eluru. Predicting traffic demand during hurricane evacuation using real-time data from transportation systems and social media. *Transportation research part C: emerging technologies*, 131:103339, 2021.
- [230] P. S. Bangare S. L. Bangare, A. Dubal and S. Patil. Reviewing otsu’s method for image thresholding. *International Journal of Applied Engineering Research*, 10(9):21777–21783, 2015.
- [231] A. Azizinamini S. Mokhtarimousavi, J. C. Anderson and M. Hadi. Improved support vector machine models for work zone crash injury severity prediction and analysis. *Transportation Research Record*, 2673(11):680–692, 2019.
- [232] S. Zhang S. Zhang, C. Zhang and J. J. Q. Yu. Attention-driven recurrent imputation for traffic speed. *IEEE Open Journal of Intelligent Transportation Systems*, 3:723–737, 2022.
- [233] A. M. Sadri, S. V. Ukkusuri, and P. Murray-Tuite. A random parameter ordered probit model to understand the mobilization time during hurricane evacuation. *Transportation Research Part C: Emerging Technologies*, 32:21–30, 2013.
- [234] Christian M Schneider, Vitaly Belik, Thomas Couronné, Zbigniew Smoreda, and Marta C González. Unravelling daily human mobility motifs. *Journal of The Royal Society Interface*, 10(84):20130246, 2013.
- [235] K. Seddig, P. Jochem, and W. Fichtner. Integrating renewable energy sources by electric vehicle fleets under uncertainty. *Energy*, 141:2145–2153, 2017.
- [236] T. Seo, A. M. Bayen, T. Kusakabe, and Y. Asakura. Traffic state estimation on highway: A comprehensive survey. *Annual reviews in control*, 43:128–151, 2017.
- [237] San Francisco Municipal Transportation Agency (SFMTA). Lifeline pass. Retrieved from <https://www.sfmta.com/>, 2022.

- [238] H. Shah, A. L. Carrel, and H. T. K. Le. Impacts of teleworking and online shopping on travel: A tour-based analysis. *Transportation*, 2022.
- [239] A. Shamshiripour, E. Rahimi, R. Shabanpour, and A. K. Mohammadian. How is covid-19 reshaping activity-travel behavior? evidence from a comprehensive survey in chicago. *Transportation Research Interdisciplinary Perspectives*, 7:100216, 2020.
- [240] S. E. Shladover. Connected and automated vehicle systems: Introduction and overview. *Journal of Intelligent Transportation Systems*, 22(3):190–200, 2018.
- [241] M. Smith and J. Castellano. Costs associated with non-residential electric vehicle supply equipment: Factors to consider in the implementation of electric vehicle charging stations. Technical report, DOE/EE-1289, 2015.
- [242] Chaoming Song, Tal Koren, Pu Wang, and Albert-László Barabási. Modelling the scaling properties of human mobility. *Nature physics*, 6(10):818–823, 2010.
- [243] Southern California Association of Governments (SCAG). 2016 Regional Travel Demand Model and Model Validation Report. Technical report, Southern California Association of Governments, 2020.
- [244] B. Staes, N. Menon, and R. L. Bertini. Analyzing transportation network performance during emergency evacuations: Evidence from hurricane irma. *Transportation research part D: transport and environment*, 95:102841, 2021.
- [245] F. Straub, S. Streppel, and D. Göhlich. Methodology for estimating the spatial and temporal power demand of private electric vehicles for an entire urban region using open data. *Energies*, 14(8):2081, 2021.
- [246] G. Tal, S. S. Raghavan, V. C. Karanam, M. P. Favetti, K. M. Sutton, J. H. Lee, and T. Turrentine. Advanced plug-in electric vehicle travel and charging behavior final report. Technical Report 12-319, California Air Resources Board Contract, 2020.

- [247] H. Tan, G. Feng, J. Feng, W. Wang, Y. J. Zhang, and F. Li. A tensor-based method for missing traffic data completion. *Transportation Research Part C: Emerging Technologies*, 28:15–27, 2013.
- [248] V. Tanksale. Anomaly detection for controller area networks using long short-term memory. *IEEE Open Journal of Intelligent Transportation Systems*, 1:253–265, 2020.
- [249] The Center for Advanced Transportation Technology (CATT) Lab. Ritis: The regional integrated transportation information system. <https://www.cattlab.umd.edu/ritis/>, 2024.
- [250] J. Tollefson. Covid curbed carbon emissions in 2020 — but not by much. *Nature*, 2021.
- [251] Transportation Secure Data Center. National renewable energy laboratory, 2017. Accessed Jan. 15, 2017: [www.nrel.gov/tsdc](http://www.nrel.gov/tsdc).
- [252] W. v Ping and K. Zhu. Evaluation of work zone capacity estimation models: A computer simulation study. In *Sixth Asia-Pacific Transportation Development Conference, 19th ICTPA Annual Meeting*, 2006.
- [253] B. Van Wee and N. Mouter. *Evaluating transport equity*, volume 7, pages 103–126. 2021.
- [254] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.
- [255] P. Vovsha, J. E. Hicks, G. Vyas, V. Livshits, K. Jeon, R. Anderson, and G. Giaimo. Combinatorial tour mode choice. Technical Report No. 17-05951, 2017.
- [256] Z. Wadud, D. MacKenzie, and P. Leiby. Help or hindrance? the travel, energy and carbon impacts of highly automated vehicles. *Transportation Research Part A: Policy and Practice*, 86:1–18, 2016.

- [257] S. Wagner, T. Brandt, and D. Neumann. Smart city planning-developing an urban charging infrastructure for electric vehicles. 2014.
- [258] X. Wan, M. C. Lucic, H. Ghazzai, and Y. Massoud. Empowering real-time traffic reporting systems with nlp-processed social media data. *IEEE Open Journal of Intelligent Transportation Systems*, 1:159–175, 2020.
- [259] D. Wang, B. Y. He, J. Gao, J. Y. Chow, K. Ozbay, and S. Iyer. Impact of covid-19 behavioral inertia on reopening strategies for new york city transit. *International Journal of Transportation Science and Technology*, 2021.
- [260] Ding Wang, Brian Yueshuai He, Jingqin Gao, Joseph YJ Chow, Kaan Ozbay, and Shri Iyer. Impact of covid-19 behavioral inertia on reopening strategies for new york city transit. *International Journal of Transportation Science and Technology*, 10(2):197–211, 2021.
- [261] Huandong Wang, Sihan Zeng, Yong Li, and Depeng Jin. Predictability and prediction of human mobility based on application-collected location data. *IEEE Transactions on Mobile Computing*, 20(7):2457–2472, 2020.
- [262] S. Wang, Z. Y. Dong, F. Luo, K. Meng, and Y. Zhang. Stochastic collaborative planning of electric vehicle charging stations and power distribution system. *IEEE Transactions on Industrial Informatics*, 14(1):321–331, 2017.
- [263] Z. Wang, X. Su, and Z. Ding. Long-term traffic prediction based on lstm encoder-decoder architecture. *IEEE Transactions on Intelligent Transportation Systems*, 22(10):6561–6571, 2020.
- [264] J. W. J. Weltevreden and O. Rotem-Mindali. Mobility effects of b2c and c2c e-commerce in the netherlands: a quantitative assessment. *Journal of Transport Geography*, 17:83–92, 2009.

- [265] X. Wen. A work zone simulation model for travel time prediction in a connected vehicle environment. arXiv preprint arXiv:1801.07579, 2018.
- [266] G. Wenzel, M. Negrete-Pincetic, D. E. Olivares, J. MacDonald, and D. S. Callaway. Real-time charging strategies for an electric vehicle aggregator to provide ancillary services. *IEEE Transactions on Smart Grid*, 9(5):5141–5151, 2017.
- [267] C. G. Wilmot. Modeling hurricane evacuation traffic: Testing the gravity and intervening opportunity models as models of destination choice in hurricane evacuation. Technical report, Louisiana Transportation Research Center, 2008. [Technical Summary] (No. LTRC technical summary report 407).
- [268] K. Winter, O. Cats, K. Martens, and B. van Arem. Relocating shared automated vehicles under parking constraints: assessing the impact of different strategies for on-street parking. *Transportation*, 48(4):1931–1965, 2021.
- [269] L. Wismans, R. van den Brink, L. Brederode, K. Zantema, and E. van Berkum. Comparison of estimation of emissions based on static and dynamic traffic assignment models. In *Annual Meeting Preprint*, pages 13–0546, 2013.
- [270] S. Wong, S. Shaheen, and J. Walker. Understanding evacuee behavior: a case study of hurricane irma, 2018.
- [271] E. W. Wood, C. L. Rames, A. Bedir, N. Crisostomo, and J. Allen. California plug-in electric vehicle infrastructure projections: 2017-2025-future infrastructure needs for reaching the state’s zero emission-vehicle deployment goals. Technical Report NREL/TP-5400-70893, National Renewable Energy Lab (NREL), Golden, CO (United States), 2018.
- [272] E. W. Wood, C. L. Rames, M. Muratori, S. Srinivasa Raghavan, and M. W. Melaina. National plug-in electric vehicle infrastructure analysis. Technical Report NREL/TP-



5400-69031; DOE/GO-102017-5040, National Renewable Energy Lab (NREL), Golden, CO (United States), 2017.

- [273] Hao Wu, Lingbo Liu, Yang Yu, Zhenghong Peng, Hongzan Jiao, and Qiang Niu. An agent-based model simulation of human mobility based on mobile phone data: How commuting relates to congestion. *ISPRS International Journal of Geo-Information*, 8(7):313, 2019.
- [274] L. Yang X. Liao, G. Wu and M. J. Barth. A real-world data-driven approach for estimating environmental impacts of traffic accidents. *Transportation Research Part D: Transport and Environment*, 117:103664, 2023.
- [275] J. Wei Y. Shen H. Qi X. Yin, G. Wu and B. Yin. Deep learning on traffic prediction: Methods, analysis, and future directions. *IEEE Transactions on Intelligent Transportation Systems*, 23(6):4927–4943, 2021.
- [276] R. Xu, Y. Guo, X. Han, X. Xia, H. Xiang, and J. Ma. Openca: an open cooperative driving automation framework integrated with co-simulation. In *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pages 1155–1162. IEEE, 2021.
- [277] P. Edara Y. Hou and C. Sun. Traffic flow forecasting for urban work zones. *IEEE Transactions on Intelligent Transportation Systems*, 16(4):1761–1770, 2014.
- [278] W. Yang, Y. Xiang, J. Liu, and C. Gu. Agent-based modeling for scale evolution of plug-in electric vehicles and charging demand. *IEEE Transactions on Power Systems*, 33(2):1915–1925, 2017.
- [279] Y. Yang, L. Mao, and S. S. Metcalf. Diffusion of hurricane evacuation behavior through a home-workplace social network: A spatially explicit agent-based simulation model. *Computers, Environment and Urban Systems*, 74:13–22, 2019.

- [280] P. M. Yellowlees, K. Chorba, M. Burke Parish, H. Wynn-Jones, and N. Nafiz. Telemedicine can make healthcare greener. *Telemedicine and E-Health*, 16(2):229–232, 2010.
- [281] B. Yu, X. Song, F. Guan, Z. Yang, and B. Yao. k-nearest neighbor model for multiple-time-step prediction of short-term traffic condition. *Journal of Transportation Engineering*, 142(6):04016018, 2016.
- [282] B. Yu, H. Yin, and Z. Zhu. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. arXiv preprint arXiv:1709.04875, 2017.
- [283] J. Yu, G. L. Chang, H. W. Ho, and Y. Liu. Variation based online travel time prediction using clustered neural networks. In *2008 11th International IEEE Conference on Intelligent Transportation Systems*, pages 85–90. IEEE, 2008.
- [284] T. Yuksel and J. J. Michalek. Effects of regional temperature on electric vehicle efficiency, range, and emissions in the united states. *Environmental Science & Technology*, 49(6):3974–3980, 2015.
- [285] Z. Hu H. Zhang J. Zhang Z. Wang, J. Guo and J. Pu. Lane transformer: A high-efficiency trajectory prediction model. *IEEE Open Journal of Intelligent Transportation Systems*, 4:2–13, 2023.
- [286] C. Zhang, J. B. Greenblatt, P. MacDougall, S. Saxena, and A. J. Prabhakar. Quantifying the benefits of electric vehicles on the future electricity grid in the midwestern united states. *Applied Energy*, 270:115174, 2020.
- [287] Jie Zhang, Baoheng Feng, Yina Wu, Pengpeng Xu, Ruimin Ke, and Ni Dong. The effect of human mobility and control measures on traffic safety during covid-19 pandemic. *PLoS one*, 16(3):e0243263, 2021.

- [288] N. Zhang and A. Alipour. A stochastic programming approach to enhance the resilience of infrastructure under weather-related risk. *Computer-Aided Civil and Infrastructure Engineering*, 38:411–432, 2023.
- [289] Fan Zhou, Yurou Dai, Qiang Gao, Pengyu Wang, and Ting Zhong. Self-supervised human mobility learning for next location prediction and trajectory classification. *Knowledge-Based Systems*, 228:107214, 2021.
- [290] P. Zhu, L. Wang, Y. Jiang, and J. Zhou. Metropolitan size and the impacts of telecommuting on personal travel. *Transportation*, 45:385–414, 2018.
- [291] Fuzhen Zhuang, Zhiyuan Qi, Keyu Duan, Dongbo Xi, Yongchun Zhu, Hengshu Zhu, Hui Xiong, and Qing He. A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1):43–76, 2020.